COMPUTATIONAL METHODS IN CONTINUUM MECHANICS

By
Edmund M. Boruh
Utah State University
Robert E. White
N.C. State University

COMPUTATIONAL METHODS IN CONTINUUM MECHANICS

By
Bolindra N. Borah
N.C. A&T State University
Robert E. White
N.C. State University

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	November 24, 1993	

4. TITLE AND SUBTITLE

NUMERICAL METHODS IN CONTINUUM MECHANICS

5. FUNDING NUMBERS

6. AUTHOR(S)

Bolindra N. Borah, Robert E. White, A. Kyrillidis,
S. Shankarlingham, Y. Ji

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Department of Mathematics
North Carolina A&T State University
Greensboro, North Carolina

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Research Office
P.O. Box 12211
Research Triangle Park, NC 27709-2211

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy, or decision, unless so designated by other documentation.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The primary objectives of this research are the development of algorithms which are applicable to non-linear continuum mechanic problems, especially to Stefan's Problems and Fluid Flows in Porous Media (Richard's Equation). We have studied 2-D and 3-D related problems. The codes are written in Cray Fortran and vectorization are done on the Cray Y-MP. The results obtained by Compact ADI, Finite Difference SOR and ADI are compared. It is seen that Compact ADI is superior.

Six scientific research papers have been presented/published in various conferences and journals in the last two years. All of the papers are attached in the appendix as attachments.

Two senior faculty members worked in this project. Bolindra N. Borah (P.I.), Professor, North Carolina A&T State University, Robert E. White (Co-P.I.), Professor, North Carolina State University are the principal researchers. Besides, four graduate students participated in this project. Two have already completed the M.S. in Applied Mathematics while the other two are completing in the Spring Semester 1994.

14. SUBJECT TERMS	15. NUMBER OF PAGES
COMPACT-ADI, FINITE DIFFERENCE SOR, STEFAN PROBLEM, INDUCED CONTRACTIVE ALGORITHM, M-MATRIX, M-FUNCTION	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

COMPUTATIONAL METHODS IN CONTINUUM MECHANICS

FINAL REPORT

Bolindra N. Borah, N.C. A&T State University
Robert E. White, N.C. State University

November 30, 1993

U. S. ARMY RESEARCH OFFICE
Research Triangle Park, N.C.

DAAL03-90-G-0216

Department of Mathematics
NORTH CAROLINA AGRICULTURAL AND TECHNICAL STATE UNIVERSITY
Greenboro, N.C. 27411

NUMERICAL METHODS IN CONTINUUM MECHANICS

FINAL TECHNICAL REPORT

TABLE OF CONTENTS

NUMERICAL METHODS IN CONTINUUM MECHANICS

FINAL TECHNICAL REPORT

1. STATEMENT OF THE PROBLEMS STUDIED

Five different problems are studied in this project. They are stated briefly below:

(a) HETEROGENEOUS DIFFUSION AND THE COMPACT VOLUME METHODS

In this paper we study heterogeneous diffusion in the context of heat conduction in laminated material, heat conduction with phase change (the Stefan problem) and fluid flow in porous media (Richards' equation). We discretize these problems via the compact volume method (CVM). We illustrate that this method is higher order accurate in the flux error than traditional methods. The resulting nonlinear algebraic systems are approximated by different versions of the SOR and ADI schemes. Vector and multiprocessing implementations were presented, which indicated a suitability of these methods for high performance computing.

(b) INDUCED CONTRACTIVE METHODS FOR FREE BVPs

In this paper we consider numerical schemes for the solution of nonlinear algebraic systems which evolve from the discretization of the Stefan problem, and from the fluid flow in a porous media problem, Richards' equation. We presented analysis of induced contractive methods for free boundary value problems. Both the ADI and SOR versions vectorize very well. In 3D problems with complicated nonlinear terms we expect the ADI version to be more robust and to perform equally as well as with the SOR version.

(c) NUMERICAL SOLUTION OF FLUID FLOW IN PARTIALLY SATURATED POROUS MEDIA

This paper describes an SOR algorithm for solving the nonlinear algebraic system which evolves from Richards' equation that models fluid

flow in a porous media. The moisture content and hydraulic conductivity functions are approximated by piecewise linear functions obtained from field data. The resulting algebraic system is solved by a variation of the nonlinear SOR algorithm. The advantage of this approach is that it avoids some of the numerical oscillations associated with large derivatives in the data. Numerical calculations are presented and illustrate the following: (i) agreement of the numerical model with observed data, (ii) dependence and comparison results as a function of uncertain data, and (iii) suitability of these algorithms for multiprocessing computations via domain decomposition methods. Extension of these algorithms to heterogeneous porous media fluid flow are discussed.

(d) A COMPACT FINITE VOLUME SCHEME FOR 2-D STEFAN PROBLEMS AND VECTOR/MULTIPROCESSOR COMPUTERS

We consider both the compact finite volume and finite difference space discretizations of the Stefan problem. The resulting algebraic systems are solved by nonlinear versions of ADI and SOR. Both algorithms contain significant parallelism which is demonstrated on two vector/multiprocessing computers, the Alliant FX/40 and the Cray Y-MP. Numerical experiments indicate that the compact discretization and ADI give the best accuracy with the minimum computational cost.

(e) A COMPARATIVE STUDY OF COMPACT FINITE VOLUME METHODS FOR THE 2-D AND 3-D DIFFUSION EQUATIONS WITH FINITE DIFFERENCE ADI AND SOR

Recently developed compact finite difference scheme (CPT) were applied to two and three dimensional diffusion equations. The relative merits of CPT-ADI were investigated with other computational schemes such as finite difference-ADI and FDM-SOR. The numerical results obtained from these three approaches are compared to known analytical solutions. According to our results CPT-ADI was found to be a superior scheme with regard to accuracy and speed.

2. SUMMARY OF THE MOST IMPORTANT RESULTS: THREE MOST IMPORTANT RESULTS ARE GIVEN BELOW:

(i) We have given an analysis of induced contractive methods for free boundary value problems. The SOR version and the ADI version of Algorithm seemed to perform equally well. However, the SOR version was very sensitive to the choice of SOR parameter, which contrasts with the ADI method which was observed to be fairly robust with respect to its acceleration parameters.

Both ADI and SOR versions vectorize very well, and as expected, in 3D problems with complicated nonlinear terms the ADI version was found to be more robust but perform equally as well as with the SOR version.

(ii) We have shown for a number of heterogeneous diffusion problems that compact volume discretization method gives higher order error estimates for both function and flux errors. Moreover, we have illustrated that both the ADI and SOR algorithms can be adapted to solve the resulting nonlinear algebraic systems. Vectorization and multiprocessing computers were shown to be effective in executing these algorithms.

In the applications to the heat conduction in a laminate materials, the Stefan problem, and Richard's equation, we have indicated how one can extend the compact volume method to 2D and 3D space problems. All these problems are of current interest to physical scientists and engineers. Moreover, the mathematical analysis of convergence problem using the more complicated versions of CVM is currently needed.

(iii) The Richard's equation is approximated by the finite difference method, and the empirical data for the moisture content and the hydraulic conductivity were approximated by pieecewise linear functions. The resulting non-linear algebraic system is solved by a variation of non-linear SOR iterative method.

Good convergence properties are observed for three types of calculations which were chosen to demonstrate the feasibility of realistic

numerical simulations using the SOR iterative methods. These includes an accurate simulation of fluid flow in Brindabella Loam, a sensitivity analysis of the computed solution upon the empirical data, and the use of multiprocessing computers via domain decomposition methods.

We expect the methods of this paper to generalize via the compact volume method to the more complicated heterogeneous case.

3. LIST OF PUBLICATIONS AND TECHNICAL REPORTS:

The following scientific articles are either published or sent for publication:

(i) Induced Contractive Methods for Free Boundary Value Problems, R. E. White, B. N. Borah, A. J. Kyrillidis; sent for publication to the Journal of Scientific Computing.

(ii) Heterogeneous Diffusion and the Compact Volume Method, R. E. White, B. N. Borah, A. J. Kyrillidis; sent for publication to Journal of Scientific Computing.

(iii) Numerical Solution of Fluid Flow in Partially Saturated Porous Media, A. J. Silva Neto, R. E. White; sent for publication to Journal of Mathematical Computing.

(iv) A Compact Finite Volume Scheme for 2-D Stefan Problems and Vector/Multiprocessor Computers, Proc. Int. Symposium, Advances in Aerospace Sciences and Engineering 1992 (Dec.), Bangalore, India, pp. 5-8.

(v) A Comparative Study of Compact Finite Volume Methods for the 2-D Diffusion Equations with Finite Difference ADI and SOR, B. N. Borah, R. E. White, A. J. Kyrillidis, S. Sankarlingham; IEEE Computer Society Proced. of the Twenty-Fourth Southeastern Symposium on System Theory, 1992, pp. 76-78.

(vi) A Comparative Study of Compact Finite Volume Methods for 3-D Diffusion Equations with Finite Difference ADI and SOR; B. N. Borah, R. E. White, A. Kyrillidis, S. Shankarlingam, Y. Ji, IEEE Computer Society, the 25th Southeastern Symposium on System Theory, 1993, pp. 1-5.

4. LIST OF PRESENTATIONS IN PROFESSIONAL MEETINGS AND CONFERENCES:

(i) Two presentations in SIAM conferences (1992, 1993).
(ii) One presentation at the International Conference on Aero-Space Science, Bangalore, India (1992).

(iii) One presentation at the 11th Army Conference on Applied Mathematics and Computing, Carnegie Mellon University, Pittsburg, PA (1993).

(iv) Three presentations in IEEE Conference (1992, 1993).

(v) One presentation at the Black College Expo, Washington, DC (1992).

5. LIST OF ALL PARTICIPATING SCIENTIFIC PERSONNEL, SHOWING ANY ADVANCED DEGREES EARNED BY THEM WHILE EMPLOYED ON THIS PROJECT:

The following senior faculty members have worked in this project:

(i) Bolindra N. Borah, Professor of Applied Mathematics (Principal Investigator), North Carolina A&T State University, Greensboro, NC.

(ii) Robert E. White, Professor of Mathematics (Co-Principal Investigator) North Carolina State University, Raleigh, NC.

The following graduate students have worked in this project:

(i) Archimedes Kyrillidis, has completed his M.S. degree in Applied Mathematics in 1991 and a Ph.D. candidate at North Carolina State University, Raleigh, NC.

(ii) Saikumar Shankarlingam, will be completing his M.S. degree in Applied Mathematics at the end of the Fall Semester, 1993.

(iii) Chardwick Sharp has completed his M.S. degree in Applied Mathematics in 1992 and is employed in industry.

(iv) Yanzhu Ji, will be completing her M.S. degree in Applied Mathematics in the Spring, 1994. She has already secured employment with an insurance company.

6. APPENDIXES

ATTACHMENT #1 Induced Contractive Methods for the Free Boundary Value Problems

Induced Contractive Methods for Free BVPs[3]

by

R. E. White[1], B. N. Borah[2], A. J. Kyrillidis[1]

Abstract. In this paper we consider numerical schemes for the solution of nonlinear algebraic systems which evolve from the discretization of the Stefan problem, and from the fluid flow in a porous media problem, Richards' equation. Both problems generate systems of the form $E + \Delta t\, A\, \beta(E) = d$ where $\beta(E)$ is continuous and nondecreasing. In the Stefan problem A will be a constant symmetric matrix. For Richards' equation A will, in general, be nonconstant and nonsymmetric. However, in both cases A is an M-matrix. We consider nonlinear SOR and nonlinear ADI schemes which may be implemented on vector/multiprocessing computers.

Subject Classifications. Primary 65H10; Secondary 76S05, 76T05.

[1] Department of Mathematics, North Carolina State University

[2] Department of Mathematics, North Carolina A&T State University

[3] Sponsored by the Army Research Office, RTP, NC, Contract No. DAAL03-90-G-0126

§1. Introduction. The primary objective of this paper is the development of algorithms which are applicable to the Stefan problem and to Richards' equation for fluid flow in a porous media, see Richards [6], Freeze and Cherry [1], and Paniconi et al. [5]. Both these problems have similar formulations, and therefore, some of the algorithms which have been used for the Stefan problem may be applicable to the more complicated Richards' equation. In Silva Neto and White [7] and [8] a local nonlinear SOR algorithm was used for the numerical solution of Richards' equation. In the present paper we describe methods which evolve from the contractive algorithm in R. E. White [11], and this is reviewed in the second section of this paper. Here we describe nonlinear SOR and nonlinear ADI iterative methods which can be used for two and three space dimension problems. We will emphasize the implementation on vector/multiprocessing computers.

In the contractive algorithm the linear solve step can be approximated by an iterative method, and an induced contractive scheme is formed. This results in a nonlinear nested iteration which is described in section three. In sections four and five we study, as special induced contractive methods, the nonlinear SOR and ADI schemes. Section six contains numerical illustrations for the Stefan problem. In section seven the similarities of the Stefan problem and of Richard's equation are discussed. Computations are done which illustrate that this method gives good comparisons with the computed and with the observed moisture data for Brindabella loam [10]. The last section contains the conclusions and recommendations.

§2. The Contractive Algorithm. In this paper we consider nonlinear problems of the form

$$E + \Delta t A \beta(E) = d \tag{1}$$

where

$$d, E \in \Re^N,$$
$$A \in \Re^{N \times N},$$
$$\beta(E) = [\beta_i(E_i)] \in \Re^N,$$
$$\beta_i : \Re \to \Re \quad i = 1, \ldots, N.$$

For the Stefan problem E is the enthalpy and $\beta(E)$ reflects the temperature. The horizontal part of Figure 1 corresponds to the latent heat.

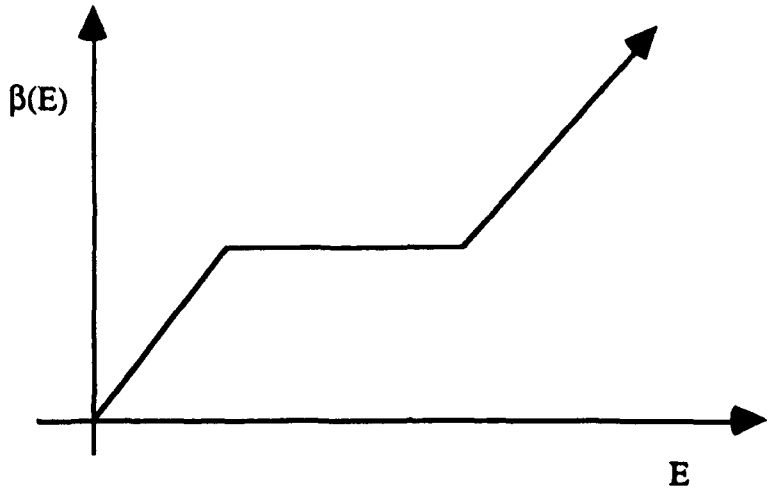


Figure 1. $\beta_i(E)$

Problems of the above form evolve from the implicit time discretization of nonlinear parabolic equations. The matrix A is from the elliptic part and is usually an M-matrix, see [4], or a symmetric positive definite (SPD) matrix. Here we emphasize the M-matrix case where A may not be symmetric. The SPD matrix case can be studied in the context of monotone mappings, see sections 6.4 and 12.1 in [4].

The function $\beta(E) = [\beta_i(E_i)] \in \Re^N$ where $E_i, \bar{E}_i > 0$ has the following properties:

$$(E_i - \bar{E}_i)\, c_1 \geq \beta_i(E_i) - \beta_i(\bar{E}_i), \qquad c_1 > 0, \tag{2.1}$$

$$\beta_i(E_i) - \beta_i(\bar{E}_i) \geq c_2\, (E_i - \bar{E}_i), \qquad c_2 \geq 0, \tag{2.2}$$

$$\beta_i(E_i) \geq c_3\, E_i\,, \qquad c_3 > 0. \tag{2.3}$$

The standard example is from the Stefan problem, but many other physical problems can be put into the above form, see [2]. Later we will emphasize the case evolving from fluid flow in porous media, see [1].

In [11] an alternate problem was proposed so that (1) could be solved via successive approximations. It has the form

$$E + \Delta t\, A(E)\, E = d \tag{3}$$

where

$$A(E) = [a_{ij}\, (\beta_j(E_j) / E_j)] \quad \text{and}$$

$$A = [a_{ij}].$$

The following algorithm was studied in [11]. There the a_{ij} where functions of E, but for the present we assume they are constants.

Algorithm 1. Contractive iteration. Consider (3). The successive approximation is

$$E^{m+1} = (I + \Delta t\ A(E^m))^{-1}\, d \equiv G(E^m).$$

If Δt is small enough, then $\rho(\Delta t\ A(E)) < \delta < 1$ for all $E > 0$, and consequently, one can write $(I + \Delta t\ A(E))^{-1}$ in the form of a geometric series. This fact and some technical arguments were used to show that G is a contraction on some box in $\Re^N$. In practice the constraint on Δt does not seem to be severe. The following is a special case of the theorem in R. E. White [11].

Theorem 1. Consider problem (3) and Algorithm 1. If A is an M-matrix in (1) and each β_i satisfies (2.1) - (2.3), then for suitably small Δt, G(E) is a contraction on some box $[\varepsilon_0, 2\|d\|_\infty]^N$.

In two and three space dimensions, the solve step will most likely be done by an iterative algorithm. Later we shall focus on the SOR and ADI algorithms.

§3. Induced Contractive Method. Let $S \subset \Re^N$ and $G: S \to S$ be contractive with $0 < c < 1$ and

$$\|G(E) - G(\bar{E})\|_\infty \le c\|E - \bar{E}\|_\infty.$$

Suppose G is as in Theorem 1 where

$$S \equiv [\varepsilon_0, 2\ \|d\|_\infty]^N \quad \text{and}$$
$$G(E) \equiv (I + \Delta t\ A(E))^{-1}\, d.$$

Algorithm 2. Induced contractive iteration. Let $I + \Delta t\, A(E) = B(E) - C(E)$ be a splitting that defines a new map $\hat{G}: S \to S$

$$\hat{G}(E) \equiv E^K$$

where

$$E^0 = E,$$
$$E^k = B(E)^{-1}\, C(E)\, E^{k-1} + B(E)^{-1}\, d \quad \text{and}$$
$$1 \le k \le K.$$

Theorem 2. Let the assumptions of Theorem 1 hold. If $\|B(E)^{-1}\, C(E)\|_\infty \le \delta < 1$ for all $E \in S$, then there exists a constant $K > 0$ such that $E^{m+1} = \hat{G}(E^m)$, where $E^0 \in S$ and $\hat{G}$ is given by Algorithm 2, converges to the solution of $E_S = G(E_S) \in S$.

Proof. Let $H = B(E)^{-1}\, C(E)$ and $\bar{H} = B(\bar{E})^{-1} C(\bar{E})$.

$$\begin{aligned}
\hat{G}(E) &= H\, E^{K-1} + B^{-1}\, d \\
&= H^K\, E + [I + \ldots + H^{K-1}]\, B^{-1}\, d \\
&= H^K\, E + (I - H^K)\, (I - H)^{-1}\, B^{-1}\, d \\
&= H^K\, E + (I - H^K)\, (B\, (I - H))^{-1}\, d \\
&= H^K\, E + (I - H^K)\, G(E) \\
&= H^K\, (E - G(E)) + G(E)
\end{aligned}$$

If $E = E_S = G(E_S)$, then $\hat{G}(E_s) = E_s$.

$$\begin{aligned}
\hat{G}(E) - \hat{G}(\bar{E}) &= H^K\, (E - G(E)) + G(E) - \bar{H}^K (\bar{E} - G(\bar{E})) - G(\bar{E}) \\
&= H^K\, (E - G(E)) - H^K (\bar{E} - G(\bar{E})) + G(E) + \\
&\qquad H^K (\bar{E} - G(\bar{E})) - \bar{H}^K (\bar{E} - G(\bar{E})) - G(\bar{E}) \\
&= H^K (E - \bar{E}) - H^K (G(E) - G(\bar{E})) + G(E) - G(\bar{E}) + (H^K - \bar{H}^K)(\bar{E} - G(\bar{E}))
\end{aligned}$$

Let $E = E^m$ with $E^{m+1} = \hat{G}(E^m)$, and $\bar{E} = E_S = G(E_S)$.

$$\begin{aligned}
E^{m+1} - E_S &= \hat{G}(E^m) - \hat{G}(E_S) \\
&= H^K\, (E^m - E_S) - H^K\, (G(E^m) - G(E_S)) + G(E^m) - G(E_S)
\end{aligned}$$

By the assumptions on H and G, we have

$$\|E^{m+1} - E_S\|_\infty \le (\delta^K + \delta^K\, c + c)\, \|E^m - E_S\|_\infty.$$

Since $\delta, c < 1$, we may choose K so that $\delta^K + \delta^K\, c + c \le (1 + c)/2 = \hat{\delta} < 1$.

Thus $\|E^{m+1} - E_S\|_\infty \le \hat{\delta}\, \|E^m - E_S\|_\infty \le \hat{\delta}^{m+1}\, \|E^0 - E_S\|_\infty$.

Corollary. Suppose $\rho(B(E)^{-1} C(E)) < 1$ for all $E \in S$. Then E^m converges to E_S provided $K = K(m)$ is suitably large.

Proof. Since $\rho(B(E)^{-1} C(E)) < 1$,

$H(E)^K = (B(E)^{-1} C(E))^K \to 0$ as $K \to \infty$.

$$\|E^{m+1} - E_S\|_\infty \le (\|H(E^m)^K\|_\infty + \|H(E^m)^K\|_\infty c + c) \|E^m - E_S\|_\infty.$$

For large enough $K = K(m)$ we have

$$\|E^{m+1} - E_S\|_\infty \le \hat{\delta} \|E^m - E_S\|_\infty \le \hat{\delta}^{m+1} \|E^0 - E_S\|_\infty, \text{ where } \hat{\delta} = (1 + c) / 2.$$

§4. Nonlinear SOR. In this section we consider the SOR version of the splitting in the induced contractive method in Algorithm 2. If $0 < \omega \le 1$, then Algorithm 2 converges to the solution of (1).

Theorem 3. Let assumptions of Theorem 1 hold. Consider the SOR splitting

$$I + \Delta t\, A(E) = \left(\frac{1}{\omega} (I + \Delta t\, D(E)) - \Delta t\, L(E)\right) - \left(\frac{1-\omega}{\omega} (I + \Delta t\, D(E)) + \Delta t\, U(E)\right)$$

Moreover, if Δt is further restricted so that $I + \Delta t\, A(E)$ is uniformly strictly diagonally dominant, then we may choose K independent of m.

Proof. $I + \Delta t\, A(E)$ is an M-matrix, and this splitting is a weak regular splitting for all $E \in S$. Thus $\rho(H(E)) < 1$, and by the above Corollary E^m must converge to E_S provided $K = K(m)$ is suitably large.

If we restrict Δt so that $I + \Delta t\, A(E)$ is, uniformly with respect to $E \in S$, strictly diagonally dominant, then we can show that the assumptions of Theorem 2 hold. This is done by standard method and is, for $\omega = 1$,

$$\|B(E)^{-1}C(E)\|_\infty \le \max_i \frac{-\Delta t \sum_{j>i} a_{ij}(E)}{1 + \Delta t a_{ii}(E) - \Delta t \sum_{j<i} a_{ij}(E)} \le \delta < 1.$$

§5. Nonlinear ADI. In this section we consider the ADI version of the induced contractive method in Algorithm 2. Here the analysis is based on M-matrices, as contrasted with SPD matrices in Ortega and Rheinboldt [4]. Later this will be useful as the analysis does not require symmetric matrices for the application to Richards' equation.

Let $A = H + V$ and $A(E) = H(E) + V(E)$ where H and V represent the grid rows and grid columns, respectively. $\hat{H} = (1/2)\, I + \Delta t\, H(E)$ and $\hat{V} = (1/2)\, I + \Delta t\, V(E)$ so that $A(E) = \hat{H} + \hat{V}$.

Theorem 4. Let the assumptions of Theorem 1 hold. Consider the ADI splitting

$$I + \Delta t\, A(E) = B(E) - C(E)$$

where

$$B(E)^{-1} = (\alpha I + \hat{V})^{-1} [I + (\alpha I - \hat{H})(\alpha I + \hat{H})^{-1}] \quad \text{and}$$
$$C(E) = A(E) - B(E).$$

There exists a constant α_0 such that if $\alpha \geq \alpha_0 \geq 0$, then Algorithm 2 converges to the solution of (1).

Proof. We shall show that the splitting is a weak regular splitting of the M-matrix $I + \Delta t\, A(E)$.

$$B(E)^{-1} = (\alpha I + \hat{V})^{-1} [I + (\alpha I - \hat{H})(\alpha I + \hat{H})^{-1}]$$
$$= (\alpha I + \hat{V})^{-1} [(\alpha I + \hat{H}) + (\alpha I - \hat{H})] (\alpha I + \hat{H})^{-1}$$
$$= (\alpha I + \hat{V})^{-1} (2\alpha I)(\alpha I + \hat{H})^{-1}.$$

By the conditions on $\beta(E)$ there exists $\alpha_1 \geq 0$ such that for $\alpha \geq \alpha_1$, $\alpha I + \hat{V}$ and $\alpha I + \hat{H}$ are M-matrices. Thus $B(E)^{-1} \geq 0$. Moreover, there exists $\alpha_2 \geq 0$ such that for $\alpha \geq \alpha_2$, $\alpha I - \hat{H}, \alpha I - \hat{V} \geq 0$. Thus for $\alpha \geq \alpha_0 = \max(\alpha_1, \alpha_2)$

$$B(E)^{-1} C(E) = (\alpha I + \hat{V})^{-1} (\alpha I - \hat{H})(\alpha I + \hat{H})^{-1} (\alpha I - \hat{V}) \geq 0.$$

By the Corollary to Theorem 2 Algorithm 2 converges to the solution of (1).

§6. Numerical Experiments for the Stefan Problem. In this section we consider the numerical solution of a one-phase Stefan problem in two space dimensions where the domain $D = (0,1) \times (0,1) \times (0,T)$, $T = 1/\sqrt{2}$. The exact solution of

$$E_t - \beta(E)_{xx} - \beta(E)_{yy} = 0$$

where

$$\beta(E) = \begin{cases} E + 1, & E < -1 \\ 0, & -1 \le E \le 0 \\ E, & E > 0 \end{cases}$$

is

$$E = \begin{cases} e^{t-(1/\sqrt{2})(x+y)} - 1, & t - (1/\sqrt{2})(x+y) \ge 0 \\ -1, & t - (1/\sqrt{2})(x+y) < 0 \end{cases}.$$

The initial condition on E and the Dirichlet boundary condition on $\beta(E)$ are implicitly given in the above formula for E. The space variables were discretized by the traditional five point finite difference method resulting in the nonlinear algebraic system (1). We compared the performance of the SOR version of the induced contractive algorithm, denoted simply as FDM-SOR, with the ADI version of the induced contractive algorithm, denoted as FDM-ADI. The computations were on the vector/multiprocessing computer Cray Y-MP. The Cray was used as a single processor.

The parameters set for this algorithm specified the number of nodes in each direction, the acceleration variables and the error tolerance. For our experiments, the number of nodes was the same for each space and time direction. Thus, the time step size is approximately forty times the maximum allowable time step size for this test problem and for the explicit method.

Both the FDM-SOR and the FDM-ADI methods showed the same pattern with respect to the number of iterations for convergence. The tolerance values were set at $\varepsilon_{in} = 10^{-5}$, $\varepsilon_{out} = 10^{-4}$, for the inner and outer iterations, respectively. The number of outer iterations was three, in all refinements considered of a basic 256 cell (16 × 16) spatial discretization, which we will subsequently refer to as the basic problem. The number of inner iterations can be considerably decreased by the correct choice of the acceleration parameter. For the SOR version of Algorithm 2 we used ω = 1.5, 1.6, 1.7, and 1.8 for the number of cells = 256, 1024, 4096, and 16,384, respectively. For the ADI version of Algorithm 2 we used α = 4.0, 6.0, 8.0, and 10.0 for the number of cells = 256, 1024, 4096, and 16,384, respectively.

Our intent was not to determine an optimal value for the acceleration parameter but rather to determine an appropriate value at which the inner iterations are minimal for our test problem. Our numerical experiments showed that the appropriate value of ω varied between 1.5 and 1.8 for all refinements of the basic problem. Our numerical experiments showed that the value of the acceleration parameter α increases considerably with successive refinements of the basic problem. Experiments showed that α can be chosen without much consideration, which is an advantage of the FDM-ADI version of Algorithm 2.

Table 1. CPU time (secs) on Cray-YMP, Algorithm 2/SOR

alg 2 \ cells	256	1024	4096
serial	0.19	2.21	27.60
vector	0.06	0.44	3.56

In our experiments we used two versions, serial and vector, of the codes that ran on the vector/multiprocessing computers. The purpose here was to determine the degree of vectorizability for each method. Tables 1 and 2 show the CPU times for both algorithms to solve each of the successive refinements of our basic mesh. In the case of FDM-SOR method the traditional red-black ordering of the nodes allows effective vectorization. In the case of FDM-ADI method we implemented the vector version of the basic tridiagonal solver.

Table 2. CPU time (secs) on Cray-YMP, Algorithm 2/ADI

alg 2 \ cells	256	1024	4096
serial	0.85	7.23	71.09
vector	0.34	2.15	16.06

Tables 1 and 2 illustrate that the FDM-SOR method exhibits higher speed-ups as compared to those of the FDM-ADI method. The CPU time of the vector FDM-SOR method was the smallest. Some of the larger CPU time for the FDM-ADI method is attributed, in this case, to the computation and storing of the components of A(E).

Next we considered a version of the ADI method that uses variable α. The calculations of the variable α were done using the scheme in [9] where $a = \Delta t$ and $b = 1/\Delta x^2$. Although this scheme was designed for a particular boundary value problem, it did work well for our nonlinear problem and was fairly robust. Table 3 shows the CPU times for the vectorized SOR, ADI with constant α, and ADI with variable α. As the number of unknowns increases, the merits of the ADI method with variable α become clear.

Table 3. CPU times for Algorithm 2

alg\cells	256	1024	4096	16 384
SOR	0.06	0.44	3.56	35.53
ADI	0.34	2.15	16.06	134.95
ADI var α	0.17	0.93	5.32	35.94

§7. Fluid Flow in a Porous Media: Richards' Equation. Richards' equation for fluid flow in a porous media was first formulated in 1931, see [6]. It is a nonlinear parabolic equation for the pressure head = Ψ where

$$h = \Psi + z,$$

and h is called the hydraulic head. The gravitational direction is given by z. Darcy's Law is used to form

$$\Theta(\Psi)_t - \nabla \cdot K(\Psi)\, \nabla h = 0 \quad \text{where}$$

K = hydraulic conductivity and Θ = moisture content. They are empirical functions of Ψ, see [1] and [3] and Figures 2 and 3

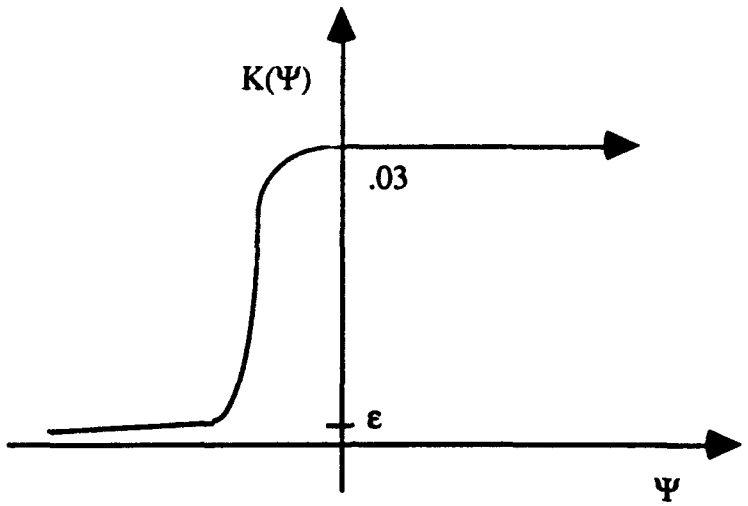


Figure 2. K = K(Ψ) = hydraulic conductivity

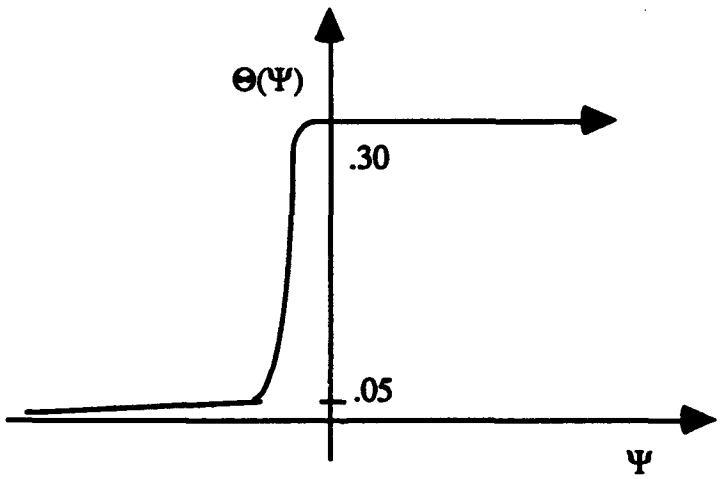


Figure 3. Θ = Θ(Ψ) = moisture content

In references [1] and [3] it is noted that the above curves may differ according to "wetting" or "drying", and the slopes can be large. The functions can be approximated by rational polynomials of order 4 or 5. This is done so that traditional analytic methods can be used; however, this introduces some approximation errors and the functions are expensive to compute.

Richards' Equation.

$$\Theta(\Psi)_t - \nabla \cdot K(\Psi)\,\nabla\Psi = K(\Psi)_z$$

Robin boundary conditions are typical and have the form $K(\Psi)\frac{dh}{dn}$ given when n is a unit outward normal to the surface.

In a recent paper by Paniconi et al. [5] several numerical methods were described, and some numerical experiments in one dimension were done. In the cases were the slopes of Θ and K were large, numerical difficulties were encountered. The moisture function $\Theta(\Psi)$ is analogous to the enthalpy function E which is a discontinuous function of temperature in the Stefan problem. This suggests that one may be able to apply the following moisture formulation of Richards' equation.

Let $E = \Theta(\Psi)$ and apply the Kirchhoff transformation to the hydraulic conductivity.

$$u = F(\Psi) \equiv \int_{\Psi_0}^{\Psi} K(\overline{\Psi})\,d\overline{\Psi},$$

$$\Psi = F^{-1}(u),$$

$$E = \Theta(F^{-1}(u)),$$

$$\beta(E) = u \text{ and}$$

$$\gamma(E) = K(\Theta^{-1}(E)).$$

Then Richards' equation can be written in terms of E where E is the primary unknown and represents the moisture.

Moisture Formulation of Richards' Equation.

$$E_t - \Delta\,\beta(E) = \gamma(E)_z$$

Here the term on the right side makes this a little more complicated than the Stefan problem. There are essentially two cases: either the derivative of γ is not too large, or the derivative is large or even a jump discontinuity occurs.

In the first case we may write $\gamma(E)_z = \gamma'(E)\ E_z$ where $0 < \gamma'(E) < M < \infty$. A reasonable implicit time discretization, say in one space dimension, is for E_i from the present time step and for $\overline{E}_i$ from the previous time step

$$\frac{E_i - \overline{E}_i}{\Delta t} + \frac{1}{\Delta x^2}(-\beta(E_{i-1}) + 2\beta(E_i) - \beta(E_{i+1})) - \frac{\gamma'(E_i)}{\Delta x}(E_{i+1} - E_i) = 0.$$

Here the coefficient matrix is more complicated than in (1), but it is still an M-matrix and the more general contractive algorithm in [11] can be used.

The second case is when $\gamma'(E)$ is so large that $\gamma(E)$ appears to have jump discontinuities. This is common once the space variables have been discretized. A special case has been studied by A. Friedman [2] where existence of a weak solution to the continuum problem is established. There for H = the Heavyside function

$$\Theta(\Psi) = \alpha\,\Psi + H(\Psi - \Psi_0),$$
$$K(\Psi) = k_1 + (k_2 - k_1)\,H(\Psi - \Psi_0),$$
$$\Psi_0 = 0,\ k_2 = 2 \text{ and } k_1 = 1.$$

Then the one space dimension form of Richards' equation is

$$(\alpha\ u + H(u))_t - u_{xx} = H(u)_x.$$

Let $E = \alpha\ u + H(u)$ and $\beta(E) = u$ so that $E = \alpha\ \beta(E) + H(u)$, $H(u) = E - \alpha\ \beta(E)$ and

$$E_t - E_x - \beta(E)_{xx} + \alpha\ \beta(E)_x = 0.$$

A reasonable discretization has the form of (1) or (3). Under some constraint on Δx A(E) will be an M-matrix, and the theorem in [11] will be applicable. In this case, A(E) is tridiagonal, and hence, there is no need to use an induced iterative method. In higher space dimensions the induced iterative methods will be more desirable.

The above approximation of the empirical functions is very restrictive. In [7] and [8] more accurate approximations are used so that diffusion of the fluid in a partially saturated medium can be calculated. A slightly less accurate approximation which will track a wet-dry interface is as follows. Here we have set $\psi_0 = 0$.

$$\theta = \begin{cases} \alpha_1 \psi + \theta_1, & \psi < 0 \\ \alpha_2 \psi + \theta_2, & \psi \geq 0 \end{cases}$$

and

$$K = \begin{cases} k_1, & \psi < 0 \\ k_2, & \psi \geq 0 \end{cases}$$

In our calculations we shifted the horizontal axis to the right so that θ and K have the graphs given in Figures 4 and 5. We used the $K(\theta)$ form of the hydraulic conductivity because it is continuous.

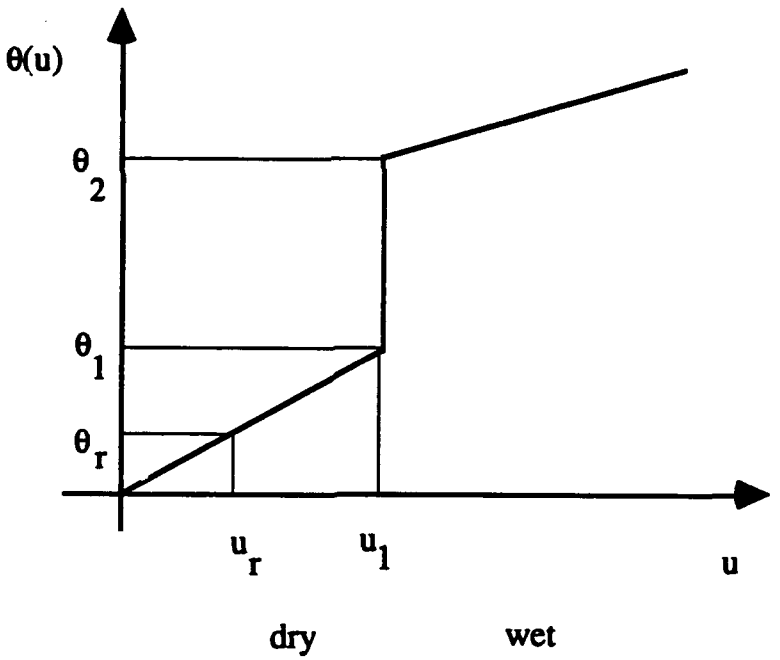


Figure 4. θ(u)

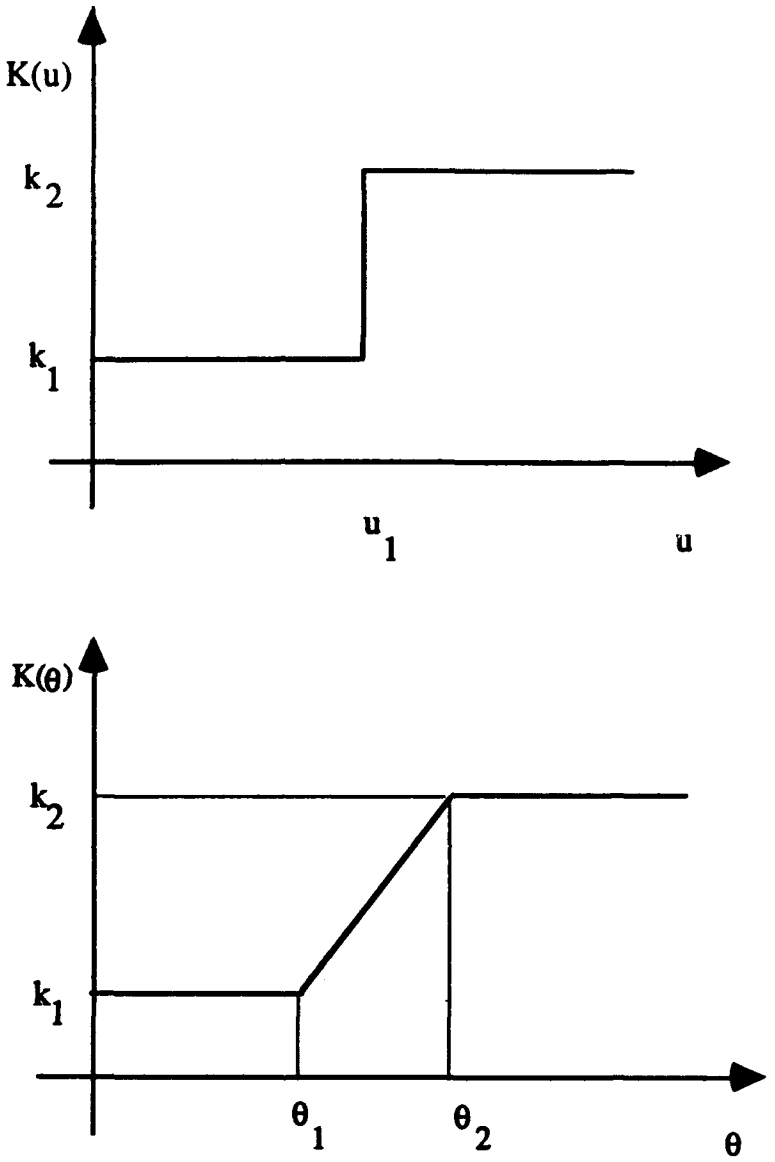


Figure 5. $K(u)$ *and* $K(\theta)$

After using the Kirchhoff transformation, Richards' equation has a moisture formulation form where

$$E = \theta,$$

$$\beta(E) = \begin{cases} \dfrac{\alpha_1}{k_1} E, & E < \theta_1 \\ \dfrac{\alpha_1}{k_1}\theta_1, & \theta_1 \le E \le \theta_2 \\ \dfrac{\alpha_2}{k_2}(E-\theta_2) + \dfrac{\alpha_1}{k_1}\theta_1, & \theta_2 < E \end{cases},$$

and

$$K(E) = \begin{cases} k_1, & E < \theta_1 \\ (k_2 - k_1)(\dfrac{E-\theta_1}{\theta_2-\theta_1}) + k_1, & \theta_1 \le E \le \theta_2 \\ k_2, & \theta_2 < E. \end{cases}$$

For the one space variable problem we have

$$E_t - \beta(E)_{xx} - K(E)_x = 0 \quad ,0 \le x \le 1, t > 0$$

with boundary conditions

$$\beta(E)_x + K(E) = R \quad ,x = 1, t > 0$$

$$\beta(E)_x + K(E) = 0 \quad ,x = 0, t > 0$$

and initial condition

$$E = \theta_r \quad ,0 \le x \le 1, t = 0.$$

The discretization for the interior cells is

$$\frac{E_i - \overline{E}_i}{\Delta t} + \frac{1}{\Delta x^2}(-\beta(E_{i-1}) + 2\beta(E_i) - \beta(E_{i+1})) - \frac{1}{\Delta x}(K(E_{i+1}) - K(E_i)) = 0.$$

The matrix in (3) is formed in a similar way, and for the interior cells

$$a_{i,i-1}(E) = -\frac{1}{\Delta x^2}\frac{\beta(E_{i-1})}{E_{i-1}},$$

$$a_{i,i}(E) = \frac{1}{\Delta t} + \frac{1}{\Delta x^2}(2\frac{\beta(E_i)}{E_i} + \Delta x \frac{K(E_i)}{E_i}) \text{ and}$$

$$a_{i,i+1}(E) = -\frac{1}{\Delta x^2}(\frac{\beta(E_{i+1})}{E_{i+1}} + \Delta x \frac{K(E_{i+1})}{E_{i+1}}).$$

The two end cells incorporate the boundary conditions in the standard way.

In the following calculations we used the parameters that reflect Brindabella loam:

$\theta_r = .11 =$ *residual moisture*	*No flow from sides or bottom*
$\theta_1 = .27 =$ *from the moisture data*	$\Delta t = .125[hr]$ *with 40 time steps*
$\theta_2 = .485 =$ *saturated moisture*	$\Delta x = .300 / 20[m]$ *with 20 grid cells*
$k_1 = 10^{-6}$	$\omega = 1.2$ *the SOR parameter*
$k_2 = .022$	$\varepsilon_{in} = 10^{-6}$ *absolute error for inner SOR*
$R = 0.0165[m / hr] =$ *flow from top*	$\varepsilon_{out} = 10^{-4}$ *absolute error for outer.*

On the average 3 or 4 inner iterations and 5 to 8 outer iterations were required for convergence. Figures 6a and 6b indicate the computed and observed moistures. In Figure 6b the diamonds, triangles and squares represent the observed data at the times 1.0, 2.25 and 5.0 hours, respectively. Note the agreement as the front progresses from the top (left) to the bottom (right). Once the bottom starts to fill the hydraulic conductivity increases and the diffusion becomes more important. In [7] the later stages of this are modelled, but the calculations are more costly than in the above model.

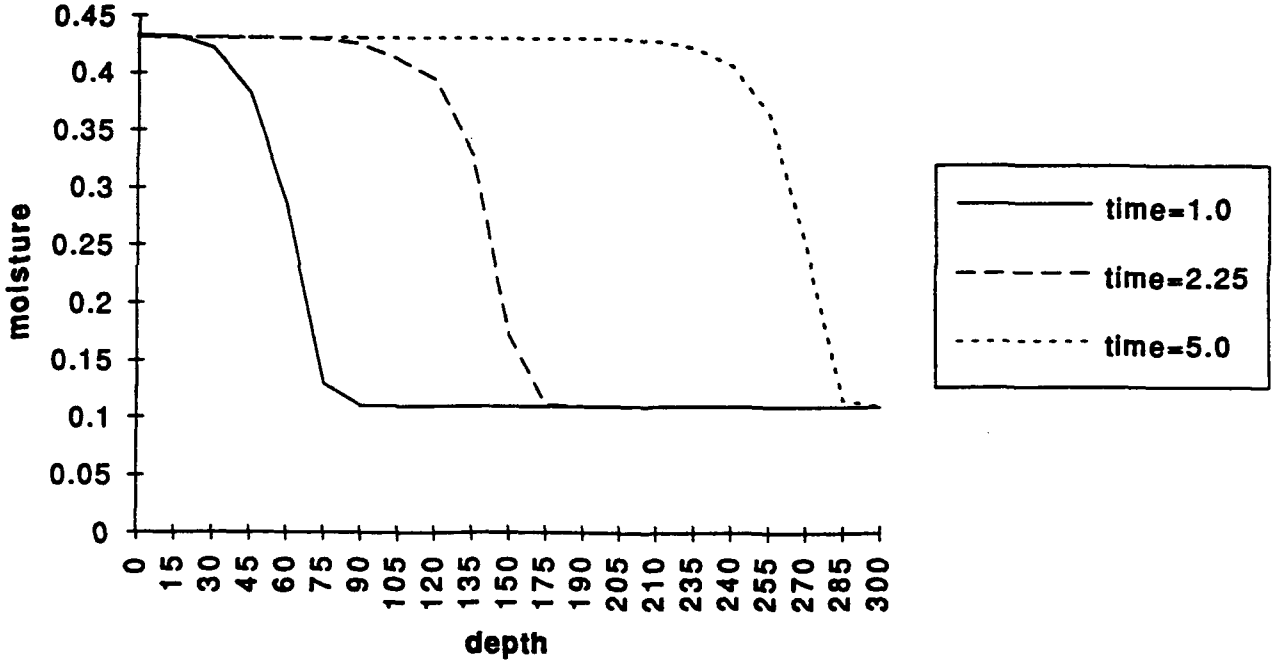


Figure 6a. Computed Moisture

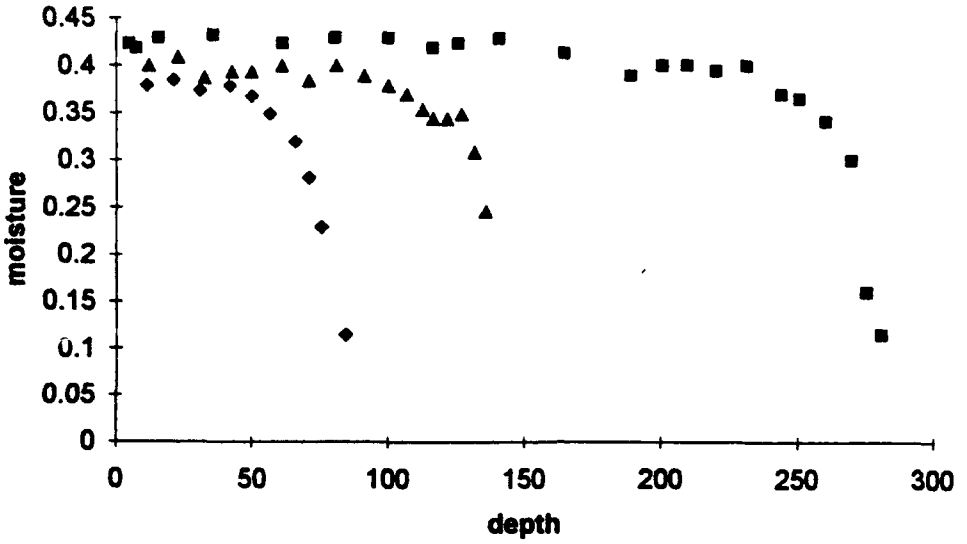


Figure 6b. Observed Moisture

§8. Conclusions. We have given an analysis of induced contractive methods for free boundary value problems. The SOR version and the ADI version of Algorithm 2 seemed to perform equally well. However, the SOR version was very sensitive to the choice of SOR parameter, and this contrasts with the ADI method which was observed to be fairly robust with respect to its acceleration parmater.

Both the ADI and SOR versions vectorize very well. In 3D problems with complicated nonlinear terms we expect the ADI version to be more robust and to perform equally well as with the SOR version.

The numerical examples were given for the two space variable Stefan problem. Application of these methods to Richards' equation was outlined. For a one space variable flow through the Brindabella loam we showed that the computed and observed moisture were in good agreement.

References.

1. R. A. Freeze and J. A. Cherry, Groundwater, Prentice-Hall Inc., Englewood Cliffs, NJ, 1979.

2. A. Friedman, "Periodic behavior for the evolutionary dam problem and related free boundary problems," in Free Boundary Problems: Application and Theory, Vol. III, edited by A. Bossavit, A. Dalmanian and M. Fremond, Research Notes in Math, 120 Pitman Advanced Publishing Program, Boston, London, Melbourne, 1983.

3. M. T. Van Genuchten and D. R. Nielsen, "On describing and predicting the hydraulic properties of unsaturated soils," Annals Geophysical, vol. 3, no. 5, 615-628, 1985.

4. J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, NY, 1970.

5. C. Paniconi, A. A. Aldama and E. F. Wood, "Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards' equation," Water Resourses, vol. 27, no. 6, 1147-1163, June 1991.

6. L. A. Richards, "Capillary conduction of liquids through porous mediums," Physics, vol. 1, 318-333, 1931.

7. A. J. Silva Neto and R. E. White, "Numerical solution of Richards' equation," to appear in the proceedings of the 12th Brazilian Congress on Mechanical Engineering, 1993.

8. A. J. Silva Neto and R. E. White, "Numerical solution of fluid flow in a partially saturated porous media," preprint 1993.

9. J. Stoer and R. Bulirsch, Introductions to Numerical Analysis, second edition, Springer-Verlag, 1993.

10. I. White and P. Broadbridge, "Constant rate rainfall infiltration: a vertical model. 2. applicationns of solutions," Water Resources Research, vol. 24, no. 1, 155-162, 1988.

11. R. E. White, "A modified finite difference scheme for the Stefan problem," Math. Comp., vol. 41, no. 164, 337-347, October 1983.

ATTACHMENT #2 Heterogeneous Diffusion and the Compact Volume Method

Heterogeneous Diffusion and the Compact Volume Method[3]

by

R. E. White[1], B. N. Borah[2], A. J. Kyrillidis[1]

Abstract. In this paper we study heterogeneous diffusion in the context of heat conduction in laminated material, heat conduction with phase change (the Stefan problem) and fluid flow in porous media (Richards' equation). We discretize these problems via the compact volume method (CVM). We illustrate that this method is higher order accurate in the flux error than traditional methods. The resulting nonlinear algebraic systems are approximated by versions of the SOR and ADI schemes. Vector and multiprocessing implementations are presented, and they indicate that these methods are suitable for high performance computing.

[1] Department of Mathematics, North Carolina State University

[2] Department of Mathematics, North Carolina A&T State University

[3] This research was sponsored by the Army Research Office, RTP, NC, Contract No. DAAL03-90-G-0126

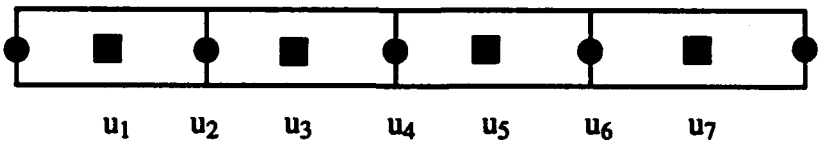
§1. Introduction. In this paper we consider diffusion problems in heterogeneous materials. We use the compact volume method (CVM), which was introduced in M. E. Rose [5], to discretize the governing differential equations. The advantage of CVM, relative to the traditional methods such as finite differences, finite elements or boundary element methods, is that CVM is higher order, in both the solution and the flux errors, accurate. Here we illustrate quadratic convergence in errors in 1D problems related to heat conduction in a laminate, heat conduction with phase change, and fluid flow in a heterogeneous porous media.

We do not give a theoretical analysis of the convergence. However, the resulting algebraic systems are more carefully studied. Some iterative methods are described, and criteria for their convergence are given. Vectorization and multiprocessing issues are discussed.

In section two we describe the CVM for linear heat conduction problems with discontinuity in the thermal conductivity. Sections three and four are applications to the Stefan problem. In section four we describe the CVM with the heat balance equation at the solid-liquid interface and develop a scheme which converges quadratically for the solid-liquid interface, the solution and the flux errors. The fifth section is an application to Richards' equation in 1D where the hydraulic conductivity is piecewise linear in the solution and has a jump discontinuity in the space variable.

§2. CVM for Linear Problems. In this section we review the general form of CVM as presented in M. E. Rose [5]. For problems of the form $-u_{xx} = f$ Rose proved quadratic convergence in both u and u_x. In order to establish the general form, consider the following example:

$$-u_{xx} = f \qquad \text{and} \qquad u(0) = 0 = u(1).$$



We partition the [0, 1] interval into 4 cells. The even nodes are inserted to insure continuity of the flux at the cell boundaries. Each cell has length equal to 2 h.

i = odd: $$-\left[\frac{u_{i+1}-u_i}{h}-\frac{u_i-u_{i-1}}{h}\right]=2\,h\,f_i \qquad \text{and}$$

i = even: $$\frac{u_{i+1}-u_i}{h}=\frac{u_i-u_{i-1}}{h}.$$

The resulting algebraic system has the following form when $u = [u_{odd}, u_{even}]^T$ where u_{odd} is the vector of odd nodes and u_{even} is the vector of even nodes.

$$\begin{bmatrix} D_1 & -C_{12} \\ -C_{21} & D_2 \end{bmatrix}\begin{bmatrix} u_{odd} \\ u_{even} \end{bmatrix}=\begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

where

$$D_1=\frac{1}{h^2}\begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}, \quad D_2=\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}$$

$$C_{12}=\frac{1}{2h^2}\begin{bmatrix} 1 & & & \\ 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \end{bmatrix}, \quad C_{21}=C_{12}^T$$

$$d_1=h^2[f_1 \quad f_2 \quad f_3 \quad f_4]^T, \quad d_2=[0 \quad 0 \quad 0]^T.$$

In this case the coefficient matrix is irreducibly diagonally dominant, an M-matrix and also symmetric positive definite. Therefore iterative methods such as SOR and ADI can be used. Moreover, when ordering the nodes as even or odd, the matrix has the form as in the red-black ordering for the FDM. Consequently, vector pipelines can be effectively used.

The coefficient matrix has a block structure which makes it easy to use block Gauss elimination. For example, in the above we may eliminate u_{odd} to get

$$[D_2 - C_{21}\,D_1^{-1}\,C_{12}]\,u_{even} = d_2 + C_{21}\,D_1^{-1}\,d_1 \qquad (2)$$

where

$$D_2 - C_{21}\,D_1^{-1}\,C_{12}=\frac{1}{4h^2}\begin{bmatrix} 2 & -1 & \\ -1 & 2 & -1 \\ & -1 & 2 \end{bmatrix}.$$

The following theorem is a well known generalization of the above.

Theorem 1. Consider the algebraic system in (1). If D_1 and $D_2 - C_{21} D_1^{-1} C_{12}$ are nonsingular, then (1) has a solution and it is given by (2) and $D_1 u_{odd} = d_1 + C_{12} u_{even}$.

The next example illustrates CVM where the media is heterogeneous, that is, the coefficient in the differential equation has a jump discontinuity with respect to the space variable.

Example 1. Heat Conduction in a 1D Laminate.

$$-(k(x)\, u_x)_x = 10\, x(1 - x),$$

$$u(0) = 0 = u(1) \qquad \text{and}$$

$$k(x) = \begin{cases} 1, & x < 0.5 \\ 2, & x \geq 0.5 \end{cases}.$$

The solution is formed by requiring u and $k\, u_x$ to be continuous at $x = 0.5$. Both the CVM and the FDM were used. The FDM has n cells and n - 1 unknowns and the CVM has n / 2 cells and n - 1 unknowns. Tables 1 and 2 illustrate, for this example, that the FDM has both function and derivative errors of order h while the CVM has both errors of order h^2.

FDM:

$$-\left[k_{i+\frac{1}{2}} \frac{u_{i+1} - u_i}{h} - k_{i-\frac{1}{2}} \frac{u_i - u_{i-1}}{h}\right] = h\, f_i$$

where $1 \leq i \leq n$ and $k_{i\pm\frac{1}{2}} = \frac{1}{2}(k(x_i) + k(x_{i\pm1}))$.

CVM:

i = odd:
$$-\left[k^-_{i+1} \frac{u_{i+1} - u_i}{h} - k^+_{i-1} \frac{u_i - u_{i-1}}{h}\right] = 2\, h\, f_i$$

where $k^{\pm}_{i\pm1} = k(x_{i\pm1} \mp \varepsilon)$, $\varepsilon > 0$ and

i = even:
$$k^+_i \frac{u_{i+1} - u_i}{h} = k^-_i \frac{u_i - u_{i-1}}{h}$$

where $k^{\pm}_i = k(x_i \pm \varepsilon)$, $\varepsilon > 0$.

Table 1. FDM for Heat Conduction in Laminate

n \ error	Fct. Error	Flux Error
8	0.002456	0.159 566
16	0.001227	0.081413
32	0.000750	0.041091
64	0.000495	0.020629
128	0.000273	0.010342

Table 2. CVM for Heat Conduction in Laminate

n \ error	Fct. Error	Flux Error
8	0.014448	0.017367
16	0.003717	0.005968
32	0.000939	0.001692
64	0.000235	0.000445
128	0.000059	0.000114

The CVM generalizes to the 2D domain in the obvious way when a rectangular grid is used. If an irregular 2D domain is required, then one may use a variation of the boundary element method (see E. K. Bruch [1]) and finite element method (see R. E. White [10]). For example, if we wish to use triangular elements and linear shape functions on $-\nabla \cdot k \nabla u = f$, then we can insert an additional center node (node o) in each element (These are analogous to the odd nodes in the 1D case), see Figure 1.

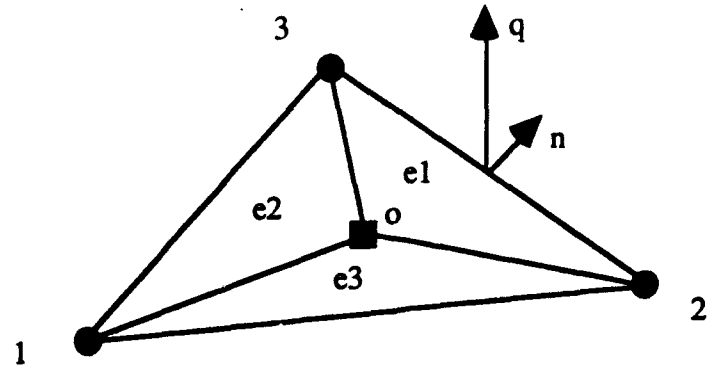


Figure 1. CVM for 2D Element

First, we apply the boundary element idea to each element and solve for u at each center node. This amounts to applying the Divergence Theorem to each element. Let $q = -k \nabla u$. Then integrate $\nabla \cdot q = f$ over e

$$\iint_e \nabla \cdot q = \iint_e f.$$

By the Divergence Theorem

$$\int_{\partial e} q \cdot n \, ds = \iint_e f.$$

Here we assume u is known at the nodes 1, 2 and 3. The u can be expressed as linear function above each subelement e_1, e_2 and e_3. Thus, the above boundary element equation has just u at the center nodes as the only unknown.

Second, we form the equations for each unknown at the vertex nodes of every element (These are analogous to the even nodes in the 1D case). This is done by requiring q to be continuous at each edge of every element. This takes the form of 3 simultaneous equations for every element, and these are easily solved. One can also combine the four equations and simultaneously solve for the four unknowns in every element. In either case the resulting system, via assembly by elements, can be solved directly or iteratively.

§3. Application to the Stefan Problem. The FDM when applied to free BVPs generates nonlinear algebraic problems of the form

$$E + \Delta t\, A\, \beta(E) = d$$

where A is a matrix associated with the elliptic part. The compact volume method, CVM, as described in [6, 7, 11] generates a larger system of the form

$$\begin{bmatrix} \frac{1}{\Delta t}E \\ 0 \end{bmatrix} + \begin{bmatrix} D_1 & -C_{12} \\ -C_{21} & D_2 \end{bmatrix} \begin{bmatrix} \beta(E) \\ u \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}. \tag{3}$$

where D_1, D_2 are diagonal matrices, C_{12} and C_{21} are rectangular matrices. $\beta(E)$ is as in the FDM and is the temperature at the center of the cells. And u is the temperature at the boundary of the cells. The diffusion equation for each cell (i = odd) is

$$\frac{E_i}{\Delta t} - \frac{1}{2h}\left[\frac{u_{i+1} - \beta(E_i)}{h} - \frac{\beta(E_i) - u_{i-1}}{h}\right] = d_i. \tag{4}$$

The flux continuity equation at cell boundary (i = even) is

$$-\frac{1}{2h}\left[\frac{\beta(E_{i+1}) - u_i}{h} - \frac{u_i - \beta(E_{i-1})}{h}\right] = 0. \tag{5}$$

Equation (3) should be viewed as a first version of the CVM method. We will want to write it in the alternate form, as in the FDM,

$$\begin{bmatrix} \frac{I}{\Delta t} + D_1(E) & -C_{12} \\ -C_{21}(E) & D_2 \end{bmatrix} \begin{bmatrix} E \\ u \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{6}$$

$$\text{where } D_1(E) = \operatorname{diag}\left[d_i \frac{\beta(E_i)}{E_i}\right] = D_1 \frac{\beta(E)}{E} \text{ and } C_{21}(E) = C_{21} \frac{\beta(E)}{E}.$$

In two and three dimensions the problem is more complicated, but it appears that ADI schemes will be useful in approximating the solution to such problems.

Consider the simplified problem (6). If D_2 is nonsingular, then we may solve for

$$u = D_2^{-1}\,[d_2 + C_{21}(E)\,E].$$

Then $$\left[\frac{I}{\Delta t} - C_{12}\,D_2^{-1}\,C_{21}(E) + D_1(E)\right]E = d_1 + C_{12}\,D_2^{-1}\,d_2 \quad \text{and}$$

$$\left[\frac{I}{\Delta t} + (D_1 - C_{12}\,D_2^{-1}\,C_{21})\,\frac{\beta(E)}{E}\right]E = d_1 + C_{12}\,D_2^{-1}\,d_2.$$

Or, $$\frac{E}{\Delta t} + (D_1 - C_{12}\,D_2^{-1}\,C_{21})\,\beta(E) = \hat{d}. \tag{7}$$

Thus, the reduced problem in (7) has the form

$$\frac{E}{\Delta t} + \hat{A}\,\beta(E) = \hat{d}$$

where $\hat{A}$ is a symmetric M-matrix. Provided $\hat{A}$ and $\hat{d}$ have been computed, SOR and ADI methods as described in section 6.4 and 12.1 in [3], will be applicable.

A more direct approach is to consider the initial problem in (6). Algorithms 1 and 2 are contractive as in the discretization given for the FDM, see [11].

Algorithm 1. Contractive.

$$\begin{bmatrix} E^{m+1} \\ u^{m+1} \end{bmatrix} = \begin{bmatrix} \frac{I}{\Delta t} + D_1(E^m) & -C_{12} \\ -C_{21}(E^m) & D_2 \end{bmatrix}^{-1} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} \tag{8}$$

Theorem 2. If $\begin{bmatrix} D_1 & -C_{12} \\ -C_{21} & D_2 \end{bmatrix}$ is an M-matrix, then for suitably small Δt, Algorithm 1 converges to the unique solution of (6).

Algorithm 2. Induced Contractive.

Let $\begin{bmatrix} \frac{I}{\Delta t}+D_1(E) & -C_{12} \\ -C_{21}(E) & D_2 \end{bmatrix} = B(E) - C(E)$ be a splitting of the matrix in (6).

Let $\begin{bmatrix} E^{m+1,0} \\ u^{m+1,0} \end{bmatrix} = \begin{bmatrix} E^m \\ u^m \end{bmatrix}$ and define the iterates for $0< k < K-1$

$$\begin{bmatrix} E^{m+1,k+1} \\ u^{m+1,k+1} \end{bmatrix} = B(E^m)^{-1}\left(C(E^m)\begin{bmatrix} E^{m+1,k} \\ u^{m+1,k} \end{bmatrix} + \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}\right) \tag{9}$$

and $\begin{bmatrix} E^{m+1} \\ u^{m+1} \end{bmatrix} = \begin{bmatrix} E^{m+1,K} \\ u^{m+1,K} \end{bmatrix}$.

The next theorem is a direct application of the results in [11].

Theorem 3. If $\rho(B(E)^{-1} C(E)) \le \delta < 1$ and the assumptions of Theorem 1 holds, then for K suitably large Algorithm 2 converges to the solution of (6).

Example 2. One phase Stefan Problem in 2D and Algorithm 2.

In this example we use Algorithm 2 to compute the numerical solution of a one-phase Stefan problem in two space dimensions where the domain $D = (0,1) \times (0,1) \times (0,T)$, $T = 1/\sqrt{2}$. The exact solution of

$$E_t - \beta(E)_{xx} - \beta(E)_{yy} = 0$$

where

$$\beta(E) = \begin{cases} E + 1, & E < -1 \\ 0, & -1 \le E \le 0 \\ E, & E > 0 \end{cases}$$

is

$$E = \begin{cases} e^{t-(1/\sqrt{2})(x+y)} - 1, & t-(1/\sqrt{2})(x+y) \ge 0 \\ -1, & t-(1/\sqrt{2})(x+y) < 0 \end{cases}.$$

The SOR splitting. Note this is vectorizable because D_1 and D_2 are diagonal matrices. The block structure in (6) reflects the red-black ordering. Vectorization is easily implemented.

The ADI splitting. Here the nodes in (6) must be reordered to reflect the classical order. For the above example the unknowns will be ordered as

$$[E_1, u_2, E_3, u_4, E_5, u_6, E_7]^T$$

This gives collections of independent tridiagonal systems, and hence, vector versions of tridiagonal solvers can be used.

This example illustrates the SOR and ADI splitting of Algorithm 2 on the above Stefan problem. We compared the performance of the SOR version of Algorithm 2, which we will denote simply as CVM-SOR, with the ADI version of Algorithm 2, denoted as CVM-ADI. The computations were done on the Cray Y-MP. The parameters set for Algorithm 2 specified the number of nodes in each direction, the acceleration variables and the tolerance. For our experiments, the number of nodes was the same for each space and time direction. Thus, the time step size is approximately forty times the maximum allowable explicit time step size for this test problem.

The tolerance values were set at $\varepsilon_{in} = 10^{-5}$, $\varepsilon_{out} = 10^{-4}$, for the inner and outer iterations, respectively. We considered four refinements of the spatial discretization, and these were for n = 16, 32, 64, and 128 cells in each direction. Our intent was not to determine an optimal value for the acceleration parameters but rather to determine an appropriate value at which the inner iterations are minimal for our test problems. For the CVM-SOR method numerical experiments showed that the appropriate values of ω were 1.5, 1.6, 1.7, and 1.8 for n = 16, 32, 64 and 128, respectively. For the CVM-ADI method numerical experiments showed appropriate values of acceleration parameter α were 4.0, 6.0, 8.0 and 10.0 for n = 16, 32, 64, and 128, respectively. Experiments showed that α can be chosen without much consideration, which is an advantage of the CVM-ADI version of Algorithm 2. We also considered a version of CVM-ADI that uses a more sophisticated acceleration scheme with variable α_i.

Table 3. CPU time (secs) on Cray Y-MP, for CVM-SOR

alg 2 \ n	n = 16	n = 32	n = 64	n = 128
serial	0.06	0.76	10.27	125.58
vector	0.02	0.12	1.14	12.11

In our experiments we used two versions, serial and vector, of the codes that ran on the Cray. The purpose here was to determine the degree of vectorizability in each method. Tables 3 and 4 show the CPU time both algorithms took to solve each of the successive refinements of our basic mesh. In the case of CVM-SOR method the block structure of the problem reflects the traditional red-black ordering of the nodes and thus allows for effective vectorization. In the case of CVM-ADI method we implemented the vector version of the basic tridiagonal solver.

Table 4. CPU time (secs) on Cray Y-MP, for CVM-ADI

alg 2 \ n	n = 16	n = 32	n = 64	n = 128
serial	0.31	2.58	27.12	254.21
vector	0.13	0.77	6.28	50.55

Clearly the CVM-SOR method exhibits higher speedups as compared to those of the CVM-ADI method. Both methods exhibit good accuracy. As noted above we also considered a version of CVM-ADI that uses variable acceleration parameters, α_i. Table 5 shows the CPU time of the vectorized CVM-SOR, CVM-ADI and the CVM-ADI with variable α_i's, denoted as ADIv. Here we see that the variable α version of CVM-ADI is competitive with CVM-SOR, and the CVM-ADIv seems to be more robust than the CVM-SOR. These results are consistent with those in [11] where the FDM is considered.

Table 5. CPU times of the Versions of Algorithm 2

alg \ n	n = 16	n = 32	n = 64	n = 128
SOR	0.02	0.12	1.14	12.11
ADI	0.13	0.77	6.28	50.55
ADIv	0.05	0.27	1.43	10.12

The single sweep ADI method also works quite well for two dimensional CVM discretizations of the form (6). Let

$$\begin{bmatrix} D_1 & -C_{12} \\ -C_{21} & D_2 \end{bmatrix} = H + V,$$

$$\hat{H} = \begin{bmatrix} \frac{I}{2\Delta t} & 0 \\ 0 & 0 \end{bmatrix} + H \quad \text{and}$$

$$\hat{V} = \begin{bmatrix} \frac{I}{2\Delta t} & 0 \\ 0 & 0 \end{bmatrix} + V.$$

$\hat{H}(E)$ and $\hat{V}(E)$ are consistent with the notation in (6). Then we can write

$$\begin{bmatrix} \frac{I}{\Delta t} + D_1(E) & -C_{12} \\ -C_{21}(E) & D_2 \end{bmatrix} = (I + \hat{H}(E)) - (I - \hat{V}(E)) = (I + \hat{V}(E)) - (I - \hat{H}(E)).$$

From here one can use the two step ADI sweeps (See Rose [6]), or implicit single ADI sweeps as follows.

Algorithm 3. Single ADI sweep.

Consider the time dependent CVM method where a sequence of algebraic problems of the form (6) are considered. Let n = time step. The single sweep ADI method is for $\hat{E} = [E, u]^T$

$$(\alpha I + \hat{H}(E^{n+1/2}))\, \hat{E}^{n+1/2} = (\alpha I - \hat{V}(E^n))\, \hat{E}^n + d \qquad \text{and} \qquad (10.1)$$

$$(\alpha I + \hat{V}(E^{n+1}))\, \hat{E}^{n+1} = (\alpha I - \hat{H}(E^{n+1/2}))\, \hat{E}^{n+1/2} + d. \qquad (10.2)$$

The nonlinear problems (10.1) and (10.2) may be approximated via the contractive

mapping; for example, in (10.1) with $\hat{d}$ the right side of (10.1)

$$\hat{E}^{k+1} = (\alpha I + \hat{H}(E^k))^{-1} \hat{d}.$$

Here α must be suitably large and can be used adaptively to obtain optimal convergence.

Example 3. One phase Stefan problem and Algorithm 3.

This example illustrates Algorithm 3 and the use of vector/multiprocessing computers. Consider the above 2D one phase Stefan problem. The computations in Table 6 were done on an Alliant FX/40 and Cray Y-MP. The Alliant FX/40 has two processors each with vector pipelines and 170 nsec clock cycle time, and the Cray Y-MP was used with one CPU, vector pipeline and 4 nsec clock.

The CVM space discretization allows effective use of independent tridiagonal solvers in either the form of vectorized tridiagonal or multiprocessing tridiagonal algorithms. In the case of the Alliant FX/40 both versions were used for the (2 cpu, vector) calculations. Table 6 records the computing times (sec) for n cells in each direction and five different computer configurations.

Table 6. ADI and CPU Times

computer \ n	n = 16	n = 32	n = 64
Alliant (serial)	2.01	15.64	122.28
Alliant (vector)	1.02	5.98	38.63
Alliant (2 cpu, vector)	0.57	3.32	21.07
Cray (serial)	0.129	0.89	6.52
Cray (vector)	0.062	0.32	1.72

§4. Application to the Stefan Problem: Higher Order Approximation.

A more complicated problem, in one dimension, is the combination of (6), (11), (12) and (13). Here we will use an nonlinear SOR scheme and use the generalization of M-matrices to M-functions as described in section 13.5 of [3]. More importantly, we will want to consider the location of the solid-liquid interface; here the flux is not continuous as in (5).

Suppose that at some iteration we observe $\beta(E_1) > u_f$ and $\beta(E_2) < u_f$. Between x_1 and x_2 there must exist an s such that $\beta(E \text{ at } s) = u_f$, see Figure 2.

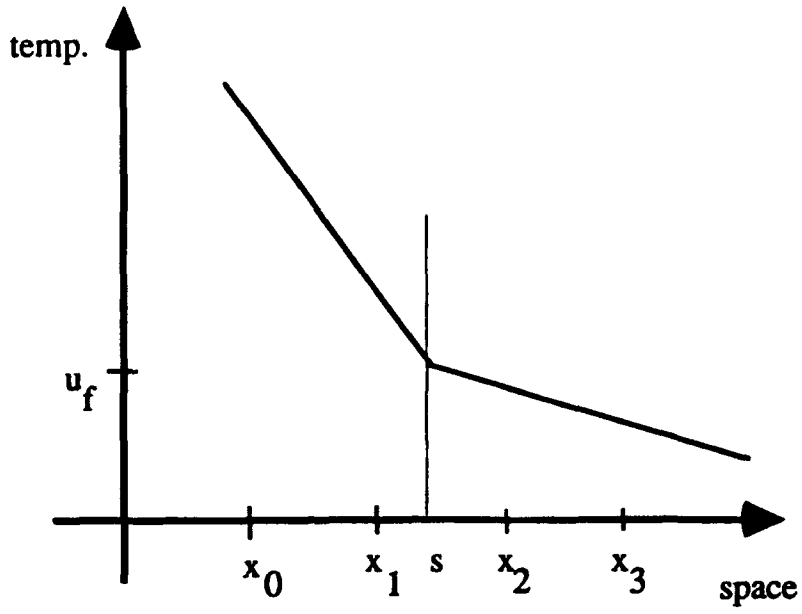


Figure 2. Phase Change

Thus (5) must be modified in three ways.
At x_1,we are in a liquid cell $[x_0, s]$ and

$$\frac{E_1}{\Delta t} - \frac{1}{s - x_0}\left[\frac{u_f - \beta(E_1)}{s - x_1} - \frac{\beta(E_1) - u_0}{h}\right] = d_1. \tag{11}$$

At s, we have the classic heat balance equation and

$$\frac{-L}{\Delta t}(s - \bar{s}) = \left[\frac{u_f - \beta(E_1)}{s - x_1} - \frac{\beta(E_2) - u_f}{x_2 - s}\right]. \tag{12}$$

At x_2, we are in a solid cell $[s, x_3]$ and

$$\frac{E_2}{\Delta t} - \frac{1}{x_3 - s}\left[\frac{u_3 - \beta(E_2)}{h} - \frac{\beta(E_2) - u_f}{x_2 - s}\right] = d_2. \tag{13}$$

Definition: The CVM for the one dimensional free BVP has the form of (6) supplemented with equations (11), (12) and (13).

The reader will note that equation (12) is a cubic with three distinct real roots. If we require

$$|s - \bar{s}| \le h, \tag{14}$$

then we must select the middle root. The condition in (14) is a discretization of $|\dot{s}|\Delta t \le h$ which has been noted as an important constraint to avoid oscillations near the solid-liquid interface, see [7].

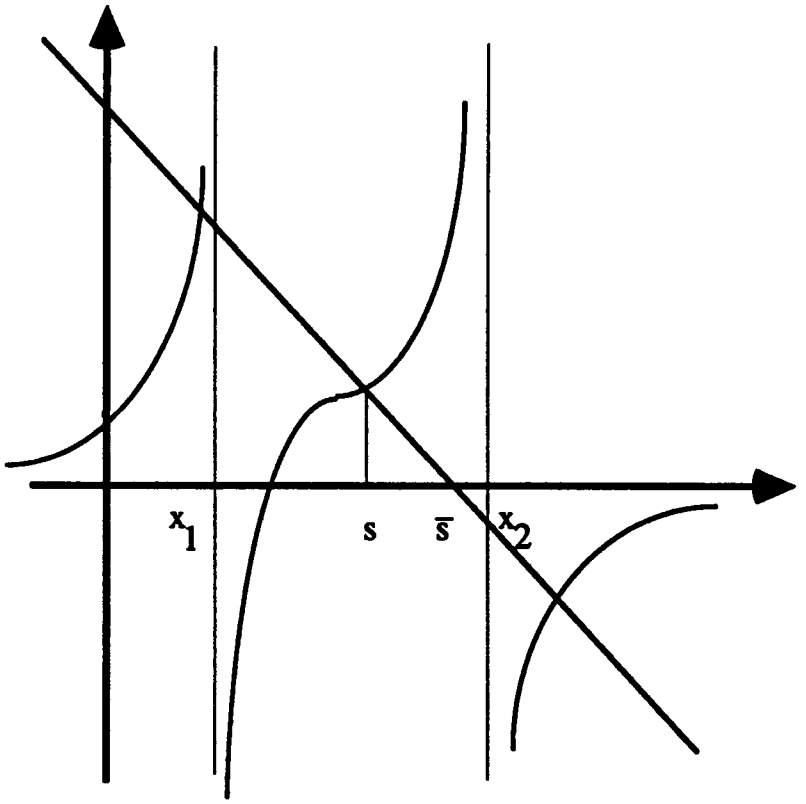


Figure 3. Solution of the Cubic Equation (12)

This seems to be a CFL type condition for the moving solid-liquid front. The analysis of (3) or (6) can be done in either the context of

$$\begin{bmatrix} D_1 & -C_{12} \\ -C_{21} & D_2 \end{bmatrix}$$

being a M-matrix or a SPD matrix. For the M-matrix case the system consisting of (6) coupled with (11), (12) and (13) is most likely an M-function, see [3]. Consequently, nonlinear SOR methods are applicable. Indeed, careful inspection of Figure 2 gives the off-diagonal antitone and the diagonally isotone conditions for the component of the system for the unknown interface. The other nodes have equations which are similar to the reduced enthalpy formulation of the Stefan problem, and the reduced system is known to be an M-functions.

Algorithm 4. Enhanced SOR for two phase Stefan problem.

Consider (6), (11), (12) and (13) where the nodes i = 1 and 2 represent a possible change in phase at some point in the iteration, require $|s - \bar{s}| \le h$ so that we must chose the center root of (12):

do an nonlinear SOR sweep of (6)
solve (12) for the center root
solve (11) and (13)
repeat until convergence.

The advantage of the more general CVM problem (6), (11), (12) and (13) is that the order of convergence is higher than (6), or the FDM. This has been observed for a number of experiments for one dimensional problems. The extension to two or three space dimensions will be a little more challenging. It appears that the solid-liquid interface condition in higher dimensions can be decoupled into equations similar to (11), (12), and (13). This suggests that a nonlinear ADI scheme can be developed. In each direction a SOR algorithm similar to Algorithm 4 can be used. Thus, we hope for two and three space dimensions, to be able to obtain higher order convergence of the discrete problem to the continuum problem, and to be able to more accurately locate the solid-liquid interface. The above schemes have a large portion of independent parts, and hence, high performance computing can be used.

Example 4. Two phase Stefan problem and solid-liquid interface.

$$E_t - \beta(E)_{xx} = 0$$

where

$$\beta(E) = \begin{cases} E, & E < 1 \\ 1, & 1 \le E \le 2 \\ (E-2)+1, & E > 2 \end{cases}.$$

The boundary and initial conditions were chosen so that the following was the solution

$$E = \begin{cases} 2\,e^{t-x}, & x < t \\ e^{t-x}, & x \geq t \end{cases}.$$

The solid-liquid interface is given by $s(t) = t$. In the calculation we started at $t = 0.2$ and ended at $t = 0.37$. Over this time interval the average function and flux errors were tabulated as well as the approximate value of the solid-liquid interface. Tables 7 illustrates quadratic (Δx^2) convergence for $\Delta t = \Delta x^2$ and this example, in the solid-liquid interface error and the function error. The convergence of the derivative error was at least linear (Δx) in all cases.

Table 7. CVM for Stefan Problem with $\Delta t = \Delta x^2$

n \ error	s Error	Fct. Error	Der. Error
10	0.003900	0.007000	0.030100
20	0.001000	0.001840	0.014100
40	0.000250	0.000510	0.005370
80	0.000060	0.000142	0.002270
160	0.000016	0.000038	0.000864

Even if the criteria on Δt is relaxed, the errors remain relatively small as is illustrated by Table 8.

Table 8. CVM for Stefan Problem with $\Delta t = 0.1\,\Delta x$

n \ error	s Error	Fct. Error	Der. Error
10	0.003900	0.007000	0.030100
20	0.001300	0.002110	0.013500
40	0.000400	0.000682	0.005270
80	0.000170	0.000247	0.002180
160	0.000075	0.000096	0.000845

§5. Application to Richards' equation. Richard's equation describes fluid flow in a porous media and was initially formulated in L. A. Richards in 1931 [4]. Presently, it is being extensively used in ground water modeling and in contamination modeling, see R. A. Freeze and J. A. Cherry [2]. This equation has strong nonlinear terms in the empirical functions for moisture and hydraulic conductivity, see [11]. Moreover, the porous media are often heterogeneous; that is, they are dependent on space position and often have discontinuities in the space variable. A simple 1D example of the steady state Richards' equation is as follows where u is the pressure and k(x,u) is the hydraulic conductivity.

Example 5. Richards' Equation in 1D.

$$-(k(x,u)\,u_x + k(x,u))_x = f(x)$$

where

$$u(0) = 0,\ u(1) = 1 + a_1 + a_0 \text{ and}$$

$$k(x,u) = \begin{cases} 3u + 0.5, & x < 0.5 \\ u + 1, & x \geq 0.5 \end{cases}.$$

The right side was chosen so that

$$u(x) = \begin{cases} x^4, & x < 0.5 \\ x^4 + a_1 x + a_0, & x \geq 0.5 \end{cases}$$

was a solution. The constants a_1 and a_0 were chosen so that u and $k\,u_x + k$ were continuous at x = 0.5. The CVM has the form

i = odd:
$$-\left[k_{i+1}\frac{u_{i+1}-u_i}{h} - k_{i-1}\frac{u_i-u_{i-1}}{h} + k_{i+1} - k_{i-1}\right] = 2hf_i$$

where $k_{i\pm1} = k(x_{i\pm1} \mp \varepsilon, u_{i\pm1}),\ \varepsilon > 0$ and

i = even:
$$k_i^+\frac{u_{i+1}-u_i}{h} + k_i^+ = k_i^-\frac{u_i-u_{i-1}}{h} + k_i^-$$

where $k_i^{\pm} = k(x \pm \varepsilon, u_i),\ \varepsilon > 0.$

The $\varepsilon > 0$ are used to insure the hydraulic conductivity is taken from the appropriate cell.

Table 9. CVM for Richards' Equation

n \ error	Fct. Error	$k\,u_x + k$ Error
8	0.05583	0.09962
16	0.01861	0.05490
32	0.00525	0.01950
64	0.00139	0.00572
128	0.00036	0.00153

In the calculations recorded in Table 9, n / 2 are the number of cells, and there are n - 1 unknowns. Nonlinear SOR was used with the hydraulic conductivity being updated as soon as any variables were computed. The function and flux = $k\,u_x + k$ errors were the max norm at the center of the cells and cell boundaries, respectively. Inspection, of the table entries indicate the CVM has quadratic convergence in both the function and flux errors.

The above example illustrates a fluid flow in a porous medium which has two layers where the pressure does not vary over a large range. Hence the hydraulic conductivity is a piecewise linear function of position and pressure. Extension to the time dependent 2D problem where the empirical functions depend only on the pressure are given in [11] and its references. These results combine to suggest that the general problem can be solved in the present context.

§6. Conclusions. We have shown for a number of heterogeneous diffusion problems that the CVM discretization method gives higher order error estimates for both the function and flux errors. Moreover, we have illustrated that both the ADI and SOR algorithms can be adapted to solve the resulting nonlinear algebraic systems. Vectorization and multiprocessing computers can be use to effectively execute these algorithms.

In the applications to the heat conduction in a laminate material, the Stefan problem, and Richards' equation we have indicated how one can extend the CVM to 2D and 3D space problems. All these problems are of current interest to physical scientists and engineers. Moreover, the mathematical analysis of convergence for the more complicated versions of CVM is certainly needed.

References

1. E. K. Bruch, The Boundary Element Method for Groundwater Flow,

2. R. A. Freeze and J. A. Cherry, Groundwater, Prentice Hall Inc., Englewood Cliffs, N. J., 1979.

3. J. M. Ortega and W. C. Rheinboldt, Iterative Solutions of Nonlinear Equations in Several Variables, Academic Press, 1970.

4. L. A. Richards, "Capillary conduction of liquids through porous mediums," Physics, vol. 1, pp. 337 -347, October 1993.

5. M. E. Rose, "Compact finite volume methods for the diffusion equations", J. Sci. Comp., vol. 4, no. 3, September 1989, pp 261 - 290.

6. M. E. Rose, "An enthalpy scheme for Stefan problems in several variables", Appl. Num. Math., vol. 12, pp 229 - 238, 1993.

7. M. E. Rose, B. N. Borah and R. E. White, "Numerical analytic studies of the Stefan problems", Report to AFSRO, June 30, 1991.

8. K. Tacke, "Discretization of the enthalpy method for planar phase change," Free Boundary Problems: Applications and Theory, IV, Pitman Advanced Publishing program, vol. 121, Boston, 1983.

9. R. E. White, "A modified finite difference scheme for the Stefan Problem," Math. Comp., vol. 41, no. 164, pp. 337 - 347, October 1983.

10. R. E. White, An Introduction to the Finite Element Method with Applications to Nonlinear Problems, Wiley Interscience, NY, 1985.

11. R. E. White, B. N. Borah and A. J. Kyrillidis, "Induced contractive methods for free BVPs", submitted 1993.

ATTACHMENT #3 Numerical Solution of Fluid Flow in Partially Saturated Porous Media

NUMERICAL SOLUTION OF FLUID FLOW IN PARTIALLY SATURATED POROUS MEDIA[1]

by

A. J. Silva Neto[2]

R. E. White[3]

Department of Mathematics
Box 8205
North Carolina State University
Raleigh, NC 27695-8205

ABSTRACT

This paper describes an SOR algorithm for solving the nonlinear algebraic system which evolves from Richards' equation that models fluid flow in a porous media. The moisture content and hydraulic conductivity functions are approximated by piecewise linear functions obtained from field data. The resulting algebraic system is solved by a variation of the nonlinear SOR algorithm. The advantage of this approach is that it avoids some of the numerical oscillations associated with large derivatives in the data. Numerical calculations are presented and illustrate the following: (i) agreement of the numerical model with observed data, (ii) dependence and comparison results as a function of uncertain data, and (iii) suitability of these algorithms for multiprocessing computations via domain decomposition methods. Extension of these algorithms to heterogeneous porous media fluid flow are discussed.

---

[1]The calculations were done at the North Carolina Supercomputing Center.
[2]Supported by CNPq and Promon Engenharia from Brazil.
[3]Supported by U. S. ARO contract number DAAL03-90-G-0126.

1. INTRODUCTION

The study of fluid flow in porous media has several important applications in engineering (Kaviany 1991 and Nield and Bejan 1992). Specific examples are: filters for industrial use, or separators in aerospace fuels (Kaviany 1991); use of geothermal energy (Rae et al. 1983 and Kimura 1989, 1989a); oil recovery (Bear 1972); groundwater (Mark 1992, 1993 and Clothier et al. 1981) and agriculture (Feng 1993).

Industrial chemical or radioactive effluents are sometimes deposited at the surface or in drums that are buried underground. In both normal operations and in accidental conditions, it is required to give an analysis of the transport of the contaminants through the soil (Muralidhar 1990, 1993). The first step in this analysis is the mathematical simulation of fluid flow through the soil.

Richards (1931) developed an equation that is a combination of the continuity equation and Darcy's law (Philip 1969), and it models fluid flow in a porous medium. It is a nonlinear parabolic partial differential equation which contains the empirical functions for *moisture content* $\theta(h)$ and *hydraulic conductivity* $K(h)$.

$$\theta(h)_t - \nabla \cdot K(h)\nabla h - K(h)_z = 0 \qquad \text{in } \Omega x(0,T) \qquad (1a)$$

where h is the hydraulic pressure head, z is the vertical direction and Ω is the space domain. The boundary condition of the third kind has the form

$$[K(h)\nabla(h+z)] \cdot n = given \qquad \text{in } Sx(0,T) \qquad (1b)$$

where n is the unit outward normal to Ω and S is the boundary of Ω. The initial condition is

$$h = \text{given} \quad \text{for } t = 0 \qquad \text{in } \Omega. \qquad (1c)$$

In general the equations (1a-1c) are coupled with a parabolic system of equations for the transport of a number of contaminants through the soil (Freeze and Cherry 1979 and Feng 1993). This is done by using the fluid velocity $v = K(h)\nabla(h + z)$ which is computed from the above system.

In practice the empirical functions for moisture content and hydraulic conductivity have several troublesome properties. First, they can have large derivatives, and this is often the case for hydraulic conductivity. Second, they are not precisely known. Third, they can have strong space dependence with jump discontinuities resulting from heterogeneous porous media. The objective of this paper is to give an approach to these problems which is based on methods used for the Stefan problem (Silva Neto and White 1993). Particular attention will be given to the first problem where there is no space dependence. In the case of space dependent empirical functions, one can use additional nodes and the continuity condition on the fluid velocity (White et al. 1993) to generalize the methods of this paper.

Traditional methods for the solution of (1a-1c) use an approximation of the empirical data by exponential functions (van Genuchten and Nielsen 1985). Then numerical methods such as Newton, Picard, or Lees implicit factored method can be used for problems without large derivatives (Paniconi et al. 1991). In addition to addressing the above problems, the approach of this paper does not involve expensive function evaluations and does eliminate the numerical oscillations associated with large derivatives of the empirical functions.

In section two we present the general approach to the problem which is adapted from the Stefan problem. Here the empirical functions are approximated by piecewise linear functions which reflect the field data (White and Broadbridge 1988). The partial differential equation is discretized by the finite difference method, and the resulting nonlinear system is solved by a nonlinear SOR algorithm (Cryer 1971) which is described in section three.

Numerical experiments are presented in sections four, five and six, and these experiments were chosen to demonstrate the feasibility of realistic two dimensional simulation of porous flows. Later, we indicate how one can extend these methods to three dimensional and heterogeneous porous flows. We show agreement of the numerical model with the field data from Brindabella silty loam soil. Also, we show how one can develop comparison results which deal with the uncertain empirical data. High performance computing issues are described. Here we demonstrate that the algorithm in section three does not vectorize well, but it does work well for multiprocessors when domain decomposition methods are used. Finally, we state our conclusions and related work.

2. DISCRETE VERSION OF RICHARDS' EQUATION

In this section we state the finite difference discretization of Richards' equation and make some comparisons with the Stefan problem. If in equation (1a) the last term is eliminated, and h were to represent temperature with K now denoting the thermal conductivity and θ the enthalpy, then this would be the enthalpy formulation of the Stefan problem (White 1985). In the Stefan problem the K and θ have jump discontinuities at the phase change temperature. SOR methods can be effectively used provided the overrelaxation is not applied during a cell's phase change.

In Richards' equation we will approximate K and θ by piecewise linear functions which could be viewed as a number of "linear phases" associated with the nonlinear flow. As in the Stefan problem we will apply the SOR method provided the cell is not changing "linear phase." Table 1 gives some data for Brindabella loam which was extrapolated from the graphs in White and Broadbridge (1988). Note, both $\theta(h)$ and K(h) are monotone, and K(h) has large derivatives.

Table 1: Brindabella Data

h [m]	θ(h) [fraction]	K(h) [m/hr]
-8.0	0.11	0.0
-1.2	0.27	0.000 118
-0.8	0.30	0.000 327
-0.7	0.31	0.000 457
-0.6	0.32	0.000 664
-0.5	0.33	0.001 138
-0.4	0.34	0.001 693
-0.3	0.36	0.002 326
-0.2	0.38	0.004 568
-0.1	0.42	0.011 117
-0.0	0.485	0.118 000

In Silva Neto and White (1993) two "linear phases" were used and were coupled with the Kirchhoff change of dependent variable. Although this was a crude approximation of the data, it did track the wet-dry interface where a rapid transition from unsaturated to saturated regions occurs. In cases where either the transition is not rapid or there is space dependence of the data, the Kirchhoff transformation is not applicable. In the following we use an implicit time discretization of Richards' equation.

$$\frac{\theta(h) - \theta(\bar{h})}{\Delta t} - (K(h)h_x)_x - (K(h)h_y)_y - K(h)_y = 0 \qquad (2)$$

where $\bar{h}$ is known from the previous time step. Here we are in two space dimensions, and y is the vertical direction.

Next we discretize the space variable by the finite difference method. In the calculations that we later discuss, we consider a two dimension flow with zero flow through the sides and bottom, and nonzero flow through the top. The finite difference grid is illustrated in Figure 1 where the nine types of boundary cells are indicated. Here there are N = 3 cells in each direction and $N^2 = 9$ unknowns.

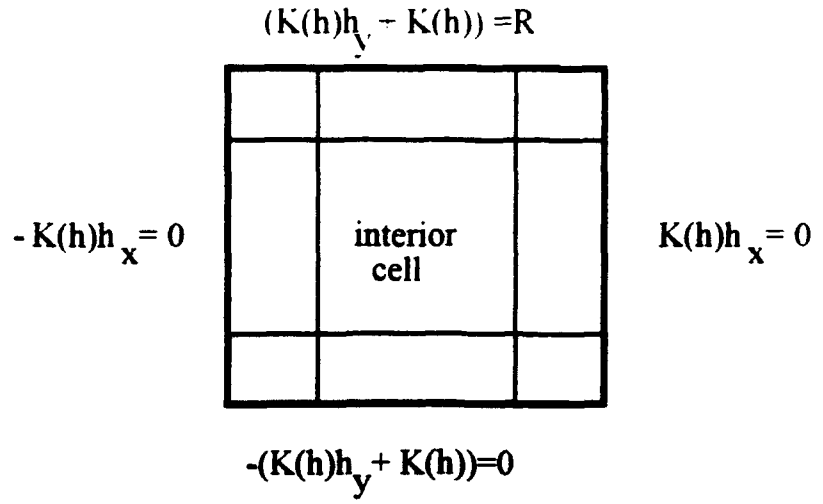


Figure 1: Finite Difference Grid

Let (i,j) denote the location in the finite difference grid. Then the general form of the finite difference equation at this location is

$$d_{i,j} - c_{i,j}h_{i,j} = \Gamma(h_{i,j}) \quad 1 \le i \le n_x \text{ and } 1 \le j \le n_y. \tag{3}$$

If (i,j) is an interior node, then

$$\Gamma(h_{i,j}) = \frac{\theta(h_{i,j})}{\Delta t} + \frac{K(h_{i,j})}{\Delta y}$$

$$c_{i,j} = (K_{i-1/2,j} + K_{i-1/2,j})\frac{1}{\Delta x^2} + (K_{i,j+1/2} + K_{i,j-1/2})\frac{1}{\Delta y^2}$$

$$d_{i,j} = \frac{\theta(\bar{h}_{i,j})}{\Delta t} + (K_{i+1/2,j}h_{i+1,j} + K_{i-1/2,j}h_{i-1,j})\frac{1}{\Delta x^2}$$
$$+ (K_{i,j+1/2}h_{i,j+1} + K_{i,j-1/2}h_{i,j-1})\frac{1}{\Delta y^2}$$
$$+ K(h_{i,j+1})\frac{1}{\Delta y}.$$

In the above equation we used the convention that $K_{i-1/2,j-1/2}$ is the average of the hydraulic conductivity at the appropriate surrounding nodes. We also will assume that the surrounding nodes are evaluated at a "previous iteration" value. Thus equation (3) is a piecewise linear system as illustrated in Figure 2. Both the piecewise linear approximation of the moisture content and hydraulic conductivity functions are monotone, and so, Γ must also be monotone nondecreasing. Since the term on the left side of equation (3) has negative slope, equation (3) has a solution, and it is unique. In Figure 2 the data is depicted as being continuous, but this need not be the case. Even if $\Gamma(h)$ has a jump and remains nondecreasing, one can still solve for a unique h (Silva Neto and White 1993).

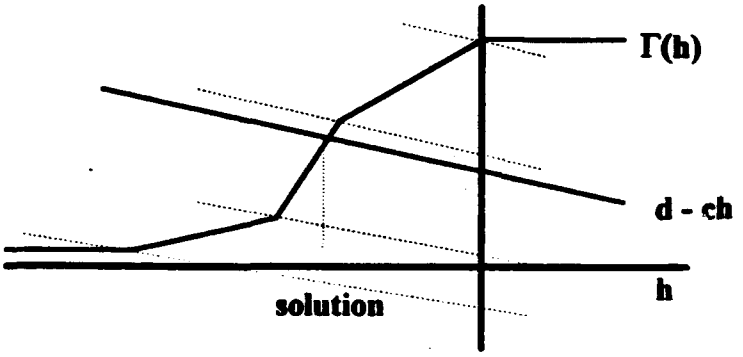


Figure 2: Solution of Equation (3)

3. NONLINEAR SOR ALGORITHM

The following algorithm has evolved from the work by Cryer (1971) for set valued systems of equations that may come from models of the Stefan problem. We apply a variation to the system given in (3). The following variables are used:

maxit = number of allowable SOR iterations per time step,

n_x, n_y = number of cells in the x and y direction,

$\overline{\omega}, \underline{\omega}$ = overrelaxation (larger than 1.0) and underrelaxation (less than 1.0) .

Nonlinear SOR Algorithm

for k = 1,maxit

for i = $1, n_x$

for j = $1, n_y$

compute $c_{i,j}$ and $d_{i,j}$ as given in (3)

solve for h in (3) as given in Figure 2

if h and $h_{i,j}$ are in the same linear phase, then

$$h_{i,j} = (1 - \overline{\omega})h_{i,j} + \overline{\omega} h$$

else

$$h_{i,j} = (1 - \underline{\omega})h_{i,j} + \underline{\omega} h$$

endif

end loop j

end loop i

test for convergence

end loop k.

In the computation of the c and d values one must consider the nine types of cells as indicated in Figure 1. For more complicated geometric configurations and for heterogeneous porous media, this will be more complicated. If adjacent cells have different moisture content and hydraulic conductivity data, then one must insert additional nodes between the cells and demand continuity of the flow velocity at the interface (White et al. 1993).

In the solve step one must determine which linear phase the solution is in, and this is done by partitioning the vertical axis as indicated in Figure 2 by the dotted lines that are parallel to the line given by $d - ch$. Hence, the solve step has a loop in it which was not indicated above. This hidden loop contains the nonlinear nature of the solve step.

Moreover, if there is a large number of linear phases, then the solve step will become more expensive to compute.

The overrelaxation is used to reduce the number of outer iterations, and the optimal choice will vary with the number of unknowns. The underrelaxation is used to avoid numerical oscillations, and this works well for choices between 0.8 and 0.9. The numerical oscillations are a result of passing from one linear phase to the next linear phase. This deals with the large derivatives of the data by breaking the changes in slopes into a number of smaller changes in slope. We found this to be much more effective than the traditional method of reducing the time step.

We experimented with a number of convergence tests. Finally, we imposed two conditions:

(i) $\max|\ \text{new } h - \text{old } h\ | \le \varepsilon_1$ and

(ii) $\iint |\ \text{new } \theta - \text{old } \theta| \le \varepsilon_2$.

The first condition is aimed at possible convergence of the pressure at each node. The second condition reflects possible convergence of the total moisture, and it is more of a global test than the first condition.

4. COMPUTATIONS FOR BRINDABELLA LOAM

The purpose of these computations is to see if our model of Richards' equation will accurately track the movement of moisture through Brindabella loam. We compare our calculations with the observations in White and Broadbridge (1988). In our numerical model we considered a 0.3[m] x 0.3[m] region with boundary conditions as indicated in Figure 1. In the top boundary we used R = 0.0165[m/hr], and the initial pressure was set as h = -8.0[m]. The moisture content function was a linear interpolation of the data in Table 1. The hydraulic conductivity data indicates a very increasing and concave up function; consequently, linear interpolation of the data would generate large errors. In the

calculations presented in Figure 3a we used a linear interpolation of the modified data for hydraulic conductivity in Table 1; we reduced the interior values by 50 percent and kept the two end values at 0.0 and 0.118.

In our computations we set the following parameters at

$\Delta t = 0.125$[hr] $\Delta x = 0.015$[m]

N = 20[cells in each dir.] R = 0.0165[m/hr]

$\omega = 0.9$ $\varpi = 1.4$

$\varepsilon_1 = 10^{-4}$ $\varepsilon_2 = 10^{-8}$.

Convergence was usually attained in 20-40 iterations. If no underrelaxation was used, then numerical oscillations would occur about once every 20 time steps.

During the initial times, the hydraulic conductivity is small, and Richards' equation is dominated by wave like properties. As time progresses the hydraulic conductivity increases so that Richards' equation is dominated by diffusion. At time 6.0[hr] the steady state solution has essentially been reached. At the bottom (y = 300[mm]) the porous media is at saturation (θ =0.485). At the top (y=0[mm]) the porous media has pressure such that K(h) = R (θ(h) =0.426). At this time the diffusion force is equal to the gravitational force; hence, no more moisture can enter the porous media.

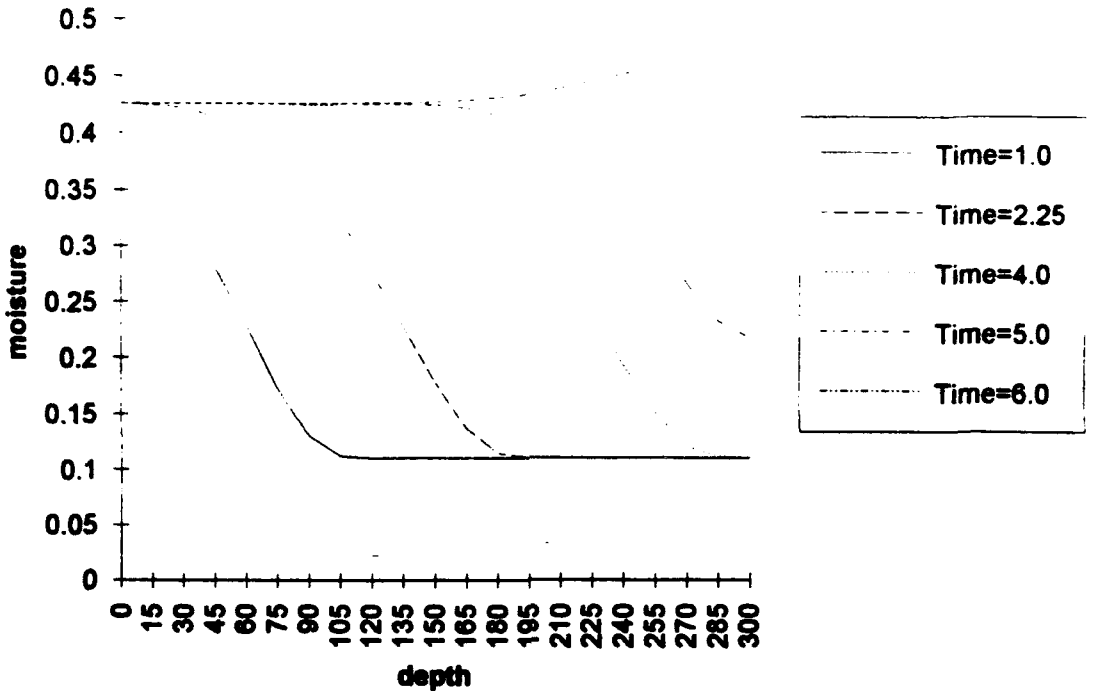


Figure 3a: Computed Moisture for Variable Times

The observed moistures are indicated by discrete points in Figure 3b. These were for the times of 1.0 (diamonds), 2.25 (triangles) and 5.0 hours (squares). The computed values give good agreement with the observed values. The computed moisture content curves are somewhat more smoothed than the observed moisture content data. This may be attributed to large hydraulic conductivity data; if one reduces the hydraulic conductivity for smaller pressures, then a sharp front can be calculated to match the observed moisture content data.

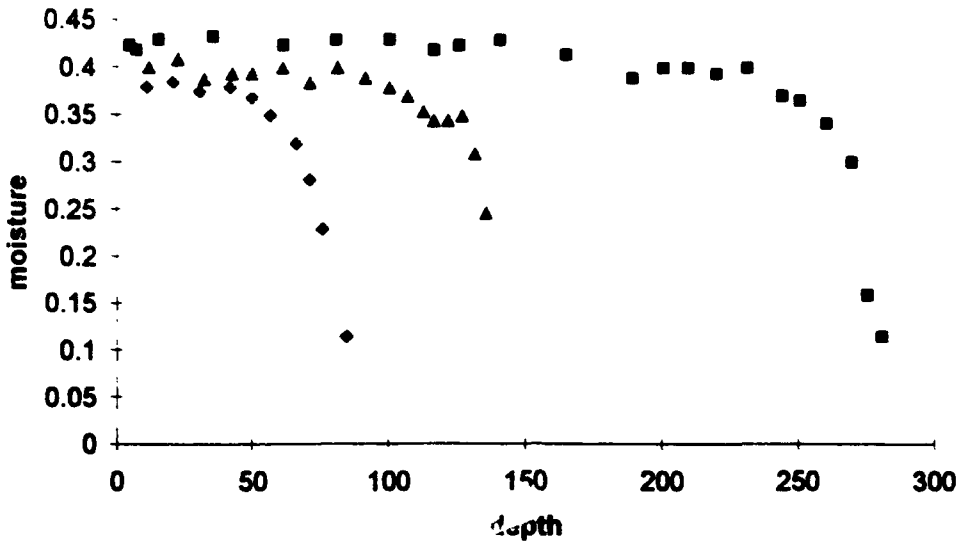


Figure 3b: Observed Moisture Data for Variable Times

5. COMPUTATIONS WITH UNCERTAIN DATA

This section contains an analysis of the moisture as a function of the empirical data of the moisture content and hydraulic conductivity functions. In practice much of this data is not precisely known, and therefore, the effects upon the computations from any model will have some uncertain aspects. In our numerical experiments we decreased K and the computed moisture at the top increased. We also increased θ and the computed moisture at the top increased. In the computations indicated in Figure 4, for time equal to 1.25[hr], we decreased K and increased θ, and the largest computed moisture content at the top was the result. Here we kept the data at the end points of Table 1 fixed and varied the interior data by increments of 20 percent. In all computations the computed moisture content at the top increased while the computed moisture content at the bottom decreased. This happens because the sides and bottom do not permit flow through them, and the total moisture must remain constant.

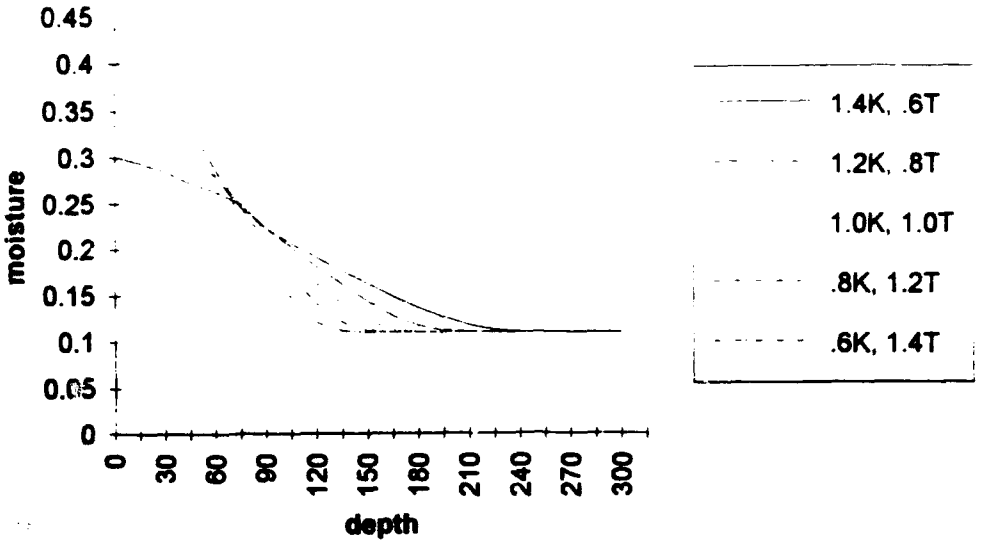


Figure 4: Moisture and Variable Data

In order to gain some insight on this, it is instructive to examine the finite difference equation at the top region. In the case of the top center nodes, d has the form

$$d = \frac{2R}{\Delta y} - \frac{K(\theta_{i,j-1})}{\Delta y} + \bar{d} \quad \text{where } \bar{d} \text{ has form similar to that given in (3).}$$

Figure 5 shows that if Γ increases, then the solution of (3) will decrease. Also, if d decreases, then the solution of (3) will decrease. Therefore, if both Γ increases and d decreases, then the solution of (3) will decrease. If K decreases, then d will increase and Γ will decrease. However, if both K decreases, and θ increases enough, then Γ will increase. For our choices of Δt and Δy this is the case. Of course, this is just an analysis at one grid point, and the argument requires much more careful discussion.

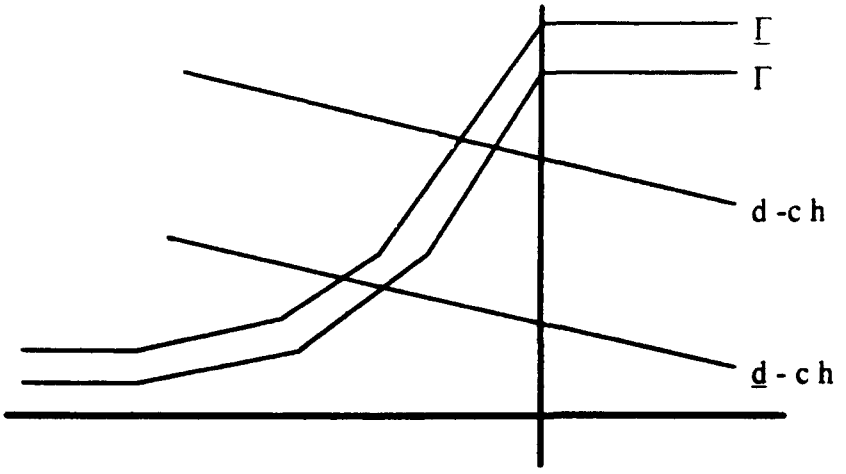


Figure 5: Moisture and Variable Data Analysis

6. COMPUTATIONS USING MULTIPROCESSORS

In the computations reported in this section we tried to implement the above algorithm on a single CPU with vectorization on a Cray Y-MP, the Alliant FX-40 with two vectorized CPUs, and the Kendall Square Research KSR1 with up to 16 CPUs and no vectorization. In the calculations in Silva Neto and White (1993) the vectorization methods did not seem to work well. These attempts involved reordering the nodes by the red-black order (checker board order). This method also did not work well for our current problem. The reason for this is that the inner most loop has computations which are too complicated to effectively be done on a vector pipeline.

The multiprocessing approach with domain decomposition reordering (White 1987. or Ortega 1988) was much more promising. This reordering is depicted in Figure 5 where L = 4 (the number of larger blocks of nodes) and the classical order of the blocks of grid points is

$$P_1, P_2, P_3, P_4, P_5, P_6, P_7 .$$

The domain decomposition order lists all the smaller interior boundary blocks first (even number blocks in Figure 6) and is

$$P_2, P_4, P_6, P_1, P_3, P_5, P_7.$$

The idea behind this reordering is to take advantage of the 5-point finite difference pattern. Once the calculations in the even blocks have been done, then the calculations in the large odd blocks are independent of one another.

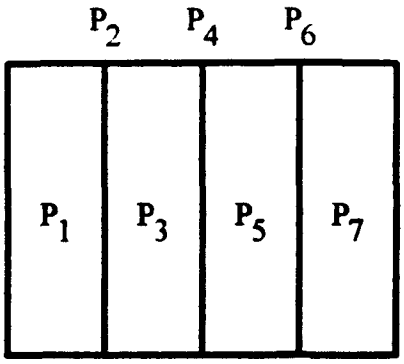


Figure 6: Domain Decomposition Order

Nonlinear SOR Algorithm: Domain Decomposition

```
for k = 1,maxit
        concurrently do SOR over the even blocks
        update
        concurrently do SOR over the odd blocks
        test for convergence
end loop k.
```

In our calculations we used the Kendall Square Research multiprocessing computer, KSR1, which is operated by the North Carolina Supercomputing Center. The KSR1 multiprocessing computer has three parallel constructs that can be used in FORTRAN code: *tile, parallel section and parallel region.* Tile is used to partition loops and is very effective for simple computations such as matrix multiplications. Parallel section can be used to concurrently execute different code segments. We used parallel region which duplicates a code segment and uses different data streams. In our

computations we controlled the number of processors by using a *team of processors* that are assigned at the beginning of the code and are used to reduce parallel overhead.

Table 2 shows the speedup and efficiency for a variety of L (the number of large blocks) and N (the number of cells in each direction). These quantities are defined as follows:

S_L = (CPU time using one block)/(CPU time using L large blocks) and

$E_L = S_L/L$.

In the first four rows N varies and L is fixed. We see increased speedup and efficiency as N increases. This is a result of decreased parallel overhead. In the last four rows N is fixed, and L is increased. Here the speedup increases, but the efficiency decreases. Of course, if there are many larger blocks (L) and the number of cells in each direction (N) remains the same, then the relative size of the larger block to the smaller block decreases. This partially accounts for the decreased efficiency. These calculations did not attempt to make the most efficient use of the FORTRAN language, or the most efficient use of the KSR1 computer's architecture.

Table 2: Speedup and Efficiency

N	L	S_L	E_L
20	2	1.56	0.78
40	2	1.68	0.84
80	2	1.75	0.88
160	2	1.78	0.89
160	4	3.16	0.79
160	8	5.09	0.64
160	16	7.29	0.46

7. CONCLUSIONS

Richards' equation was approximated by the finite difference method, and the empirical data for the moisture content and the hydraulic conductivity were approximated by piecewise linear functions. The resulting nonlinear algebraic system was solved by a variation of the nonlinear SOR iterative method. Good convergence properties were observed for three types of calculations which were chosen to demonstrate the feasibility of realistic numerical simulations using this method. These included an accurate simulation of fluid flow in Brindabella loam, a sensitivity analysis of the computed solution upon the empirical data, and the use of multiprocessing computers via domain decomposition methods.

In the above calculations the empirical data did not have a space dependence. However, in White et al. (1993) we illustrated for a steady state and one space dimension version of Richards' equation that the compact volume method, in place of the finite difference method, could be effectively used for such heterogeneous problems. The compact volume method can be viewed as an enhanced finite difference method where additional nodes are inserted at the cell interface and additional equations are generated by requiring continuity of the fluid velocity at these interfaces. This may be done for all cell interfaces or for just those cells where the empirical functions change with respect to the space variable. We expect the methods of this paper to generalize via the compact volume method to the more complicated heterogeneous case.

REFERENCES

Bear, J. (1972), *Dynamics of Fluids in Porous Media*, Dover, New York.

Clothier, B.E., White, I., and Hamilton, G.J. (1981), "Constant rate rainfall infiltration: field experiments," *Soil Sci. Soc. Am. J., vol. 45, pp. 245-249.*

Cryer, C.W. (1971), "The solution of a quadratic programing problem using systematic overrelaxation," *SIAM J. Control*, vol. 9, no. 3. pp.385-392.

Feng, J., (1993), *Modeling of Chemical Transport from Agricultural Waste Lagoons*, Ph.D. Dissertation, North Carolina State University.

Freeze, R.A. and Cherry, J.A. (1979), *Groundwater*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey.

Kaviany, M. (1991), *Principles of Heat Transfer in Porous Media*, Mechanical Engineering Series, Springer-Verlag, New York.

Kimura, S. (1989), "Transient forced convection heat transfer from a circular cylinder in a saturated porous medium," *Int. J. Heat and Mass Transfer*, vol. 32, no. 1, pp.192-195.

Kimura, S. (1989a), "Transient forced natural convection heat transfer about a vertical cylinder in a porous medium," *Int. J. Heat and Mass Tranfer*, vol.32, no. 3, pp.617-620.

Mark, S.M., (1992), "Groundwater modeling: a vital tool for water resource management," *Colorado School of Mines Quarterly Review*, vol. 92, no. 4, pp.9-12.

Mark, S.M., (1993), "Groundwater modeling: a vital tool for water resource management, Part II," *Colorado School of Mines Quarterly Review*, vol. 93, no. 1, pp.1-5.

Muralidhar, K. (1990), "Flow and transport in single rock fractures," *J. Fluid Mech.*, vol. 215, pp.481-502.

Muralidhar, K. (1993), "Near-field solution for heat and mass transfer from buried nuclear waste canisters," *Int. J. Heat and Mass Transfer*, vol. 36, no. 10, pp.2665-2674.

Nield, D.A. and Bejan, A. (1992), *Convection in Porous Media*, Springer-Verlag, New York.

Ortega, J.M. (1988), *Introduction to Parallel and Vector Solution of Linear Systems*, Plenum Press, New York.

Paniconi, C., Aldama, A.A. and Wood, E.F. (1991), "Numerical evaluation of iterative and noniterative methods for the solution of the nonlinear Richards' equation," *Water Resources Research*, vol. 27, no. 6, pp.1147-1163.

Philip, J.R. (1969), "Theory of infiltration," *Adv. Hydrosci.*, vol. 5, pp.215-296.

Rae, J., Robinson, P.C. and Wickens, L.M. (1983), "Coupled heat and groundwater flow in porous rock," *Numerical Methods in Heat Transfer, Volume II*, edited by Lewis, R.W., Morgan, K. and Schrefler, B.A., Wiley, New York.

Richards, L.A. (1931), "Capillary conduction of liquids through porous media," *Physics*, vol.1, pp.318-333.

Silva Neto, J.A. and White, R.E. (1993), "Numerical solution of Richards' equation," to appear in the proceedings of the 12th Brazilian Congress of Mechanical Engineering.

van Genuchten, M.T. and Nielsen, D.R. (1985), "On describing and predicting the hydraulic properties of unsaturated soils," *Ann. Geophys.*, vol. 3, no. 5, pp.615-628.

White, R.E. (1985), *An Introduction to the Finite Element Method with Applications to Nonlinear Problems*, Wiley, New York.

White, R.E. (1987), "Multisplitting and parallel iterative methods," *J. Comp. Meth. in Appl. Mech. and Eng.*, vol. 64, pp.567-577.

White, I. and Broadbridge, P. (1988), "Constant rate rainfall infiltration: a vertical model. 2. Applications of solutions," *Water Resouces Research*, vol. 24. no.1, pp.155-162.

White, R.E., Borah, B.N. and Kyrillidis, A.J. (1993), "Heterogeneous diffusion and the compact volume method," submitted.

ATTACHMENT #4 A Compact Finite Volume Scheme for 2-D Stefan Problems and Vector/Multiprocessor Computers

Recent Advances in Aerospace Sciences and Engineering

Proceedings of the International Symposium

Volume I

Editors

H S Mukunda
A V Krishna Murty

INTERLINE
PUBLISHING
Bangalore, India.

A Compact Finite Volume Scheme for 2–D Stefan Problems and Vector/Multiprocessor Computers*

R.E. WHITE[1], B.N. BORAH[2] and A.J. KYRILLIDIS[2]

Abstract

We consider both the compact finite volume and finite difference space discretizations of the Stefan problem. The resulting algebraic systems are solved by nonlinear versions of ADI and SOR. Both algorithms contain significant parallelism which is demonstrated on two vector/multiprocessing computers, the Alliant FX/40 and the Cray Y–MP. Numerical experiments indicate that the compact discretization and ADI give the best accuracy with the minimum computational cost.

Introduction

This report outlines some new numerical methods for the solution of multiple space dimension Stefan problem. This will include both SOR and ADI algorithms for the nonlinear algebraic systems which result from both the FDM and compact space discretizations methos. Analysis of convergence will be discused.

The ADI algorithms have been introduced because they appear, in many examples, to be more effective than some of the traditional algorithms related to the Stefan problem. Moreover, these ADI algorithms, as in the linear parabolic problems, have large number of independent tridiagonal solvers. Therefore they can be effectively implemented on vector/multiprocessing computers. This is illustrated for Alliant FX/40 and the Cray Y–MP.

In this paper, we will use the enthalpy formulation of the Stefan problem. (See Rose)[3], Elliot[2] and White[5]).

1. $E_t - \Delta\,\beta(E) = f$ on $D \times (O, T)$

   $E(x,O)$ = given on D

   $\beta(E(x,t))$ = given on $\partial D \times (O, T)$

where

$\beta(E)$=Kirchoff transformed temperature

$\beta(E) = u$ is the "inverse" of the enthalpy function $E = H(u)$.

Typical values are (See Williams and Wilson[7])

$u_f = 270$

$a_1 = 1.OE+2$

$a_2 = 1.OE+3$ (Degrees in Kelvin)

L=15

Stefan Number = $1 = (u - u_f)/L$

For these typical values and the range of u, it is important to note that

$\beta(E)/E$ is Lipschitz continuous and

$O \le m \le (\beta(E))/E \le M$ where

$m = \beta\ (27000))/27000 \approx 270/27000 = 1.0E{-}2$

$M = \beta\ (27030))/27030 \approx 285/27030 = 1.054E{-}2$

The difference M–m = 5.4E–4 is relatively small and will be important in a mild constraint on the time interval = Δt.

Problem (1) may be discretized in the time variable either explicitly (see Athey[1]) or implicitly (see White[5]). In the former, this is a stability condition on Δt, but there is not a nonlinear system to solve. The implicit time discretization yields no stability constraint on ΔT, but it does give a nonlinear system to solve. In White[5] and Elliot[2] nonlinear SOR methods have been used to approximate this system which has the form

2. $E + \Delta t\, A\, \beta\ (E) = d$

Where A is an M–matrix associated with $-\Delta$.

Later in White[6], a modified version of (2) was studied so that more traditional linear solvers could be used in the solution of (2). The component from of (2) is

$$E_i + \Delta t \sum_{j=i}^{n} a_{i,j}\, \beta\,(E_j) = d_i$$

3. $$E_i + \Delta t \sum_{j=i}^{n} \left(a_{i,j} \frac{\beta\ (E_j)}{E_j} \right) E_j = d_i$$

Let $A(E) = \left[a_{i,j} \dfrac{\beta\ (e_j)}{E_j} \right]$ and consider the vector form of (3)

$E + \Delta t\, A(E)\, E = d$

4. $I + \Delta t\, A(E))\, E = d$

*This research was sponsored by the Air Force Office of Scientific Research, Bolling Air Force Base, D C Contract No. F49620-89-C-0010 and Army Research Office, Research Triangle Park, N C Contract No.DAAL03-90-G-0126.

1. NC State University, Raleigh, NC
2. NC A & T State University, Greensboro, NC

For Δt somewhat restricted and the above typical values, $I+\Delta t\, A(E)$ will be strictly diagonally dominant, and therefore, nonsingular. This allows us to define the following algorithm:

5. $E^{m+1} = (I + \Delta t\ A(E^m))^{-1}\ d$

The convergence of (5) was established by showing for suitably small Δt that the map was

$$G(E) \equiv (I + \Delta t\, y_o A(E^m))^{-1}\, d$$

contractive. However, numerical experiments on a one variable Stefan problem showed that Δt constraint note to be too severe.

This study extends the solution of[1] and [4] to higher dimensions by considering ADI splitting associated with A(E).

ADI: One H and V sweep per time step

Let A = H + V, where H and V are associated with horizontal grid rows and the vertical grid columns of - Δ. Then we define

$$\bar{H}(E) \equiv H(E) = \left[h_{i,j}\,\frac{\beta(E_j)}{E_j}\right]$$

$$\bar{V}(E) \equiv V(E) = \left[v_{i,j}\,\frac{\beta(E_j)}{E_j}\right]$$

$$\hat{\Lambda}(E) = \frac{I}{2} + \Delta t\,\bar{H}(E)$$

$$\hat{\Diamond}(E) = \frac{I}{2} + \Delta t\,\bar{V}(E)$$

Note $I + \Delta t\ A(E) = \hat{\Lambda}(E) + \hat{\Diamond}(E)$, and for mild constraint on Δt, $\hat{\Lambda}(E)$ and $\hat{\Diamond}(E)$ are nonsingular and triadiagonal.

This suggests the following nonlinear verion of the Peaceman – Ranchford ADI algorithm for[1]:

6.1

$$\frac{E^{k+1/2} - E^k}{\Delta t/2} + \bar{H}(E^{k+1/2})\,E^{k+1/2} = f^{k+1/2} - \bar{V}(E^k)$$

6.2

$$\frac{E^{k+1} - E^{k+1/2}}{\Delta t/2} + \bar{V}(E^{k+1})\,E^{k+1} = f^{k+1} - \bar{H}(E^{k+1/2})$$

At each time step, k, (6.1) and (6.2) gives a nonlinear system to solve:

7.1 $\hat{\Lambda}(E^{k+1/2})\,E^{k+1/2} = d^{k+1/2} - \hat{\Diamond}(E^k)\,E^k$

7.2 $\hat{\Diamond}(E^{k+1})\,E^{k+1} = d^{k+1} - \hat{\Lambda}(E^{k+1/2})\,E^{k+1/2}$

It is useful to incorporate the ADI acceleration parameter α which will also alow as to avoid any further constraint on Δt:

$$A(E) = (\alpha I + \hat{\Lambda}(E)) - (\alpha I - \hat{\Diamond}(E))$$

Then (7.1) and (7.2) give

8.1

$$(\alpha I + \hat{\Lambda}(E^{k+1/2}))\,E^{k+1/2} = d^{k+1/2} + (\alpha I - \hat{\Diamond}(E^{k+1/2}))E^k$$

8.2

$$(\alpha I + \hat{\Diamond}(E^{k+1}))\,E^{k+1} = d^{k+1} + (\alpha I - \hat{\Lambda}(E^{k+1/2}))\,E^{k+1/2}$$

Both (8.1) and (8.2) can be solved as in (5). For example, consider (8.1) with k + 1/2 suppressed and m equals the iteration index.

Consider

9 $(\alpha I + \hat{\Lambda}(E)\,d$

10 $E^{m+1} = (\alpha I + \hat{\Lambda}(E^m))^{-1}\,d = \left(I + \frac{1}{\alpha}\hat{\Lambda}(E^m)\right)^{-1} d$

Proposition 1:

Let A, $A(E)$, $\hat{\Lambda}(E)$, $\hat{\Diamond}(E)$ be as above. If A is an M-matrix and $\beta(E)$ is as defined above, then A(E), $\hat{\Lambda}(E)$, $\hat{\Diamond}(E)$ are non-singular for small Δt. Moreover, there exists $\alpha_o > 0$ such that for $\alpha \geq \alpha_0$[10] converges to the unique solution of [9].

The proof follows that given in White[6] for[4] and [5].

In practice α is used as an acceleration parameter and usually only 3–5 iterations are required to solve (8.1) or (8.2). The schemes in (8.1) and (8.2) do not solve the implicit time step:

11. $\frac{E^{k+1} - E^k}{\Delta t} + A(E^{k+1})\,E^{k+1} = f^{k+1}$

For each fixed problem in[11] there are a number of ADI schemes with multiple H and V sweeps. We focus on only one in the next section.

ADI: Multiple H and V sweeps per time stop

Consider the simple nonlinerar system[4] with

$A(E)\,E = \bar{H}(E) + \bar{V}(E)$, or

12. $(\hat{\Lambda}(E) + \hat{\Diamond}(E))\,E = d$

After linearization the multiple H and V sweep ADI algorithm is

13.1

$$(\alpha I + \hat{\Lambda}(E^{m_0}))\,E^{m+1/2} = d + (\alpha I - \hat{\Diamond}(E^{m_0}))\,E^m$$

13.2

$$(\alpha I + \hat{\Diamond}(E^{m_0}))\,E^{m+1} = d + (\alpha I - \hat{\Lambda}(E^{m_0}))\,E^{m+1/2}$$

One can solve (13.1) and insert it into (13.2) to obtain

14. $E^{m+1} = \delta(E^{m_0}) + H(E^{m_0})\,E^m$

If E^{m+1} converges, define E^{m_0+1}

provided the spectral radius of H (E^{m_0}) is less than 1. This is

15. $E^{m_0+1} = (I - H(E^{m_0}))^{-1}\,\delta(E^{m_0})$

ine (15) forms the outer iteration of our algorithm to solve (2).

Proposition 2

Consider (12) and the algorithms (13) - (15). Let the same assumptions hold as in proposition 1. Then there exists a $\bar{\alpha}_o > 0$ such that if $\alpha \geq \bar{\alpha}_o$, the inner iterations (14) converge to E^{mo+1} in (15), and E^{mo+1} converges to the solution in (12).

The proof requires the spectral radius of H(E) to be uniformly bounded below 1.0. This step makes use of monotonicity properties (nonnegative matrices), and it contrasts with the linear case where symmetric positive definite matrices are used.

The above schemes is not the only way to do multiple H and V sweeps, but it does allow one to give a convergence analysis. The best version of multiple H and V sweeps is not yet clear. The numerical experiments that we have done indicate that the Peaceman-Rachford method (one H and V sweep per time step) is adequate.

Numerical Experiments

In this section we consider the numerical solution of

$$E_t - \Delta \beta(E) = 0$$

where the domain $D = (0,1) \times (0,1) \times (0,1/\sqrt{2})$

and

$$E = \begin{cases} e^{t - \frac{1}{\sqrt{2}}(x+y)} - 1, & t - \frac{1}{\sqrt{2}}(x+y) \geq 0 \\ -1, & t - \frac{1}{\sqrt{2}}(x+y) < 0 \end{cases}$$

The space variables are discretized in two ways, the compact finite volume (CPT) (See Rose[4]), and the traditional five-point finite difference method (FDM). The resulting algebraic problems are solved by ADI (one H and V sweep per time step) and by SOR. The computations were done on two vector/multiprocessing computers: the Alliant FX/40 and the Cray Y-MP. The Alliant FX/40 has two processors each with a vector pipeline. The Cray Y-MP was used as a single processor with a more sofisticated vector pipeline and much shorter clock cycle time (4 nsec as compared to 170 nsec).

The CPT space discretization has a great deal of parallelism which can be executed on either vector pipelines or multiprocessing architectures. The unknowns at the center of the cells maybe grouped together, and the unknowns at the boundary of the cells grouped together. The resulting coefficient matrix has the form of a two coloring scheme. Hence one can execute the SOR scheme for the CPT discretization on vector pipelines. Also the CPT space discretization when solved by ADI methods will have a large number of independent tridiagonal solvers. Thus, the vector version of the tridiagonal algorithm may be used.

In the case of FDM space discretization, the traditional red-black ordering of nodes allows vectorization of the SOR algorithm. Also, domain-decomposition methods have been used on SOR so that the two processors of the Alliant FX/40 can be effectively used. Of course, we can use the vector version of the tridiagonal algorithm to solve the FDM when ADI is used.

Table 1 CPT–ADI

Cells	16^2	32^2	64^2
Computers			
Alliant, serial	2.01	15.64	122.28
Alliant, vector	1.02	5.98	38.63
Alliant, vector, 2 processors	0.57	3.32	21.07
Cray, serial	0.129	0.89	6.52
Cray, vector	0.062	0.32	1.76

Table 2 CPT–SOR

Cells	16^2	32^2	64^2
Computers			
Alliant, serial	4.68	74.44	1155.00
Alliant, vector	2.67	30.72	508.81
Alliant, vector, 2 processors	1.41	15.99	270.91
Cray, serial	0.41	6.23	96.24
Cray, vector	0.11	1.06	11.34

Table 1 contains computing time done for the compact discretization and using one sweep ADI algorithm. Both computers were effective vectorizers. Table 2 contains computing times done for the compact discretization with red-black ordering SOR algorithm where ω=1.2

Table 3 FDM–SOR

Cells	16^2	32^2	64^2
Computers			
Alliant, serial	0.82	12.93	186.72
Alliant, vector	0.89	10.29	124.97
Alliant, vector, 2 processors	0.51	5.58	67.35
Cray, serial	0.069	1.05	15.48
Cray, vector	0.043	0.42	4.19

The third table contains computing times for the FDM with *red- black ordering for SOR where* ω=1.2. This method requires more iterations to reach convergence than does the CPT- SOR method.

All three methods give good accuracy. The CPT-ADI seems to be more accurate and less computing time required for this and related examples.

References

1. D.R. Athey, Epidemiology "A finite difference scheme for melting problems," J.Inst. Math Appl. Vol. 13 (1974), pp. 353–366.

2. C.M. Elliot, "On the finite element approximation of elliptic variational inequality arising from an implicit time discretization of the Stefan problem.. RMA J. Num. Analy. Vol. I (1981), pp. 115–125

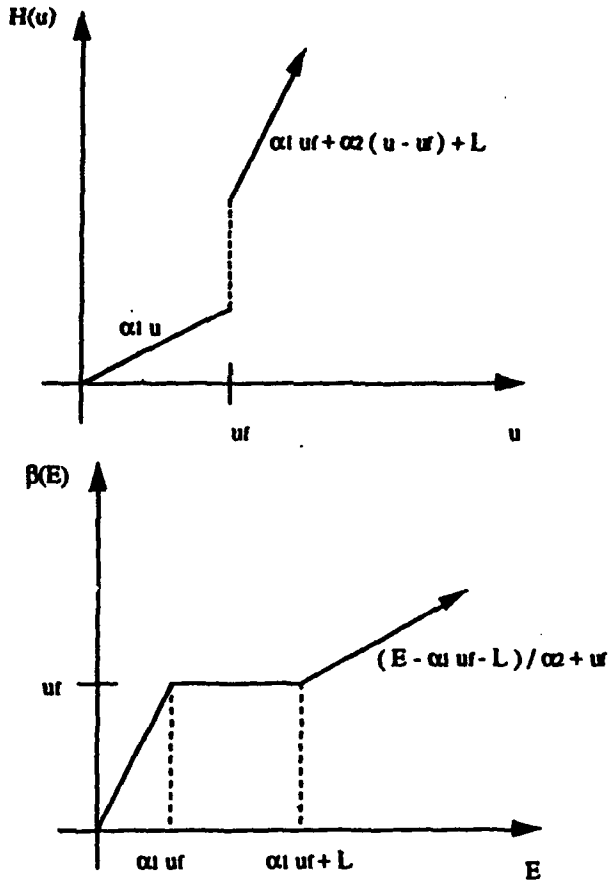
3. M.E. Rose, "A method for calculating solutions of parabolic equations with free boundary," Math. Comp. Vol. 14 (1960), pp. 249–256.

4. M.E. Rose, "Compact finite volume methods for the diffusion equations," J.Sci.Comp. Vol 4. No.3 (1989), pp. 261–290.

5. R.E. White, "A numerical solution of the enthalpy formulation of the Stefan problem," SIAM J. of Num. Analy Vol. 19 (1982), pp. 1158–1173.

6. R.E. White, "A modified finite difference scheme for the Stefan problem." Math. of Comp. Vol. 41, No.164 (1983), pp. 337–347.

7. M.A. Williams and D.G. Wilson, "Iterative solution of a nonlinear system arising in phase-change problems," SIAM J. Sci. Comp., Vol. 11, No. 6 (1990), pp. 1087–1101.



Typical values are (See Williams and Wilson [7])

$u_f = 270$

$a_1 = 1.0E+2$

$a_2 = 1.0E+3$ (Degrees in Kelvin)

$L = 15$

Stefan Number = $1 = (u - u_f)/L$

For these typical values and the range of u, it is important to note that

$(\beta(E))/E$ is Lipschitz continuous and

$0 \le m \le (\beta(E))/E \le M$ where

$m = (\beta(27000))/27000 = 270/27000 = 1.0E\text{-}2$

$M = (\beta(27030))/27030 = 285/27030 = 1.054E\text{-}2$

The difference M - m = 5.4E-4 is relatively small and will be important in a mild constraint on the time interval = Δt.

ATTACHMENT #5 A Comparative Study of Compact Finite Volume Methods for 2-D Diffusion Equations with Finite Difference ADI and SOR

A COMPARATIVE STUDY OF COMPACT FINITE VOLUME METHODS FOR THE 2-D DIFUSSION EQUATION WITH FINITE DIFFERENCE ADI AND SOR

B.N. Borah
R.E. White
A.J. Kyrillidis
S. Sankarlingham

Reprinted from PROCEEDINGS OF THE TWENTY-FOURTH SOUTHEASTERN SYMPOSIUM ON SYSTEM THEORY, THE THIRD ANNUAL SYMPOSIUM ON COMMUNICATIONS, SIGNAL PROCESSING EXPERT SYSTEMS, AND ASIC VLSI DESIGN (SSST/CSA '92), Greensboro, North Carolina

A Comparative Study of Compact Finite Volume Methods for the 2-D Difussion Equation with Finite Difference ADI and SOR*

B. N. Borah[1], R. E. White[2], A. J. Kyrillidis[1], S. Sankarlingham[1]

1. NC State University, Raleigh, NC
2. NC A&T State University, Greensboro, NC

Abstract

Recently developed Compact Finite Difference scheme (CPT) is applied to two dimensional diffusion equations. The relative merits of CPT-ADI are investigated with other computational schemes such as finite difference method - ADI (FDM-ADI) and FDM-SOR. The numerical results obtained from these three approaches are compared to known analytical solutions. The primary interest of this study lies on vectorization and parallel processing. According to our results shown in tables 1 CPT-ADI is found to be superior scheme with regards to accuracy, than both FDM-ADI and FDM-SOR. It is also fastest algorithm than both FDM-ADI and FDM-SOR as it is evident from CPU times.

1. Introduction

We shall describe briefly the compact finite difference method (CPT) for one dimmensional steady state problem. The extension to 2-D problem may be easily done. The underlying physics behind this approach lies in solving the differential equation in isolation from its neighboring subintervals (i.e. compactly). Then, extend the solution in the large by means of continuity conditions for the flux and temperature accross the boundaries of the contiguous subintervals (See Rose [1]).

We consider here one dimensional steady diffusion problem:

(1) D. E. $u' = f$
(2) $f' = u$
(3) B. C. $f = g$
for $X \in I$ and $I = [I_-, I+]$

Divide the interval I into m nonoverlaping subintervals:

$$I_j = \left\{ x \mid x_{j-\frac{1}{2}} \le x \le x_{j+\frac{1}{2}} \right\},$$

with center points x_j, j=1/2,3/2,..., M - 1/2 and interior endpoints x_j, j=1,2,..., M -1. We shall adopt the finite difference notations $\Delta x_i = x_{i+1/2} - x_{i-1/2}$. $h_i = \Delta x/2$ and u(i) $=u_i$. We also denote:

$I_c = \{1/2, 3/2,..., M - 1/2\}$ (center points)
$I_e = \{1,2,..., M - 1\}$ (interior endpoints)

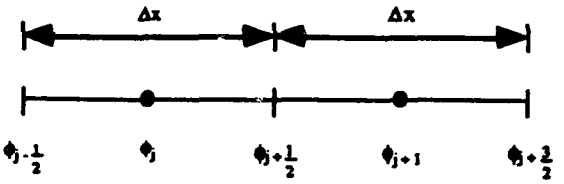


Fig. 1.

The discrete equations for (1) and (2) can be written as, for $j \in I_c$,

(4) $$u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}} = \Delta x_j f_j$$

(5a) $$\phi_{j+\frac{1}{2}} - \phi_j = h_j u_{j+\frac{1}{2}}$$

and

(5b) $$\phi_j - \phi_{j-\frac{1}{2}} = h_j u_{j-\frac{1}{2}}$$

Next, it is required that both u and f be continuous across every endpoint common to two intervals:

$$[u]_i = [\phi]_i = 0 \quad \text{for } i \in I_e.$$

* This research was sponsored by the Air Force Office of Scientific Research, Bolling Air Force Base, D. C. Contract No. F49620-89-C-0010 and Army Research Office, Research Triangle Park, N. C. Contract No. DAAL03-90-G-0126.

0-8186-2665-8/92 $03.00 © 1992 IEEE

Using this continuity condition, $[u]_i = 0$ in terms of ϕ we have

$$\frac{\phi_i - \phi_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} = \frac{\phi_{i+\frac{1}{2}} - \phi_i}{h_{i+\frac{1}{2}}}$$

and in the case of equal intervals, which reduces to

$$\phi_i = \frac{\phi_{i+\frac{1}{2}} + \phi_{i-\frac{1}{2}}}{2}, \quad i \in I_e \text{ (endpoints)} \tag{6}$$

Now from (4) and (5) we get,

$$\phi_{j+\frac{1}{2}} - 2\phi_j + \phi_{j-\frac{1}{2}} = 2h_j^2 f_j, \quad i \in I_c \tag{7a}$$

and from (6)

$$\phi_{j+\frac{1}{2}} - 2\phi_j + \phi_{j-\frac{1}{2}} = 0, \quad i \in I_e \tag{7b}$$

The equations (7) and the boundary conditions lead to determined system of algebraic equations for the values of ϕ. These equations lead to a tridiagonal system of equations which is, therefore, solved by Thomas algorithm.

The extension to two dimensional scheme may be easily done. Two dimensional stencil is shown in Fig. 2.

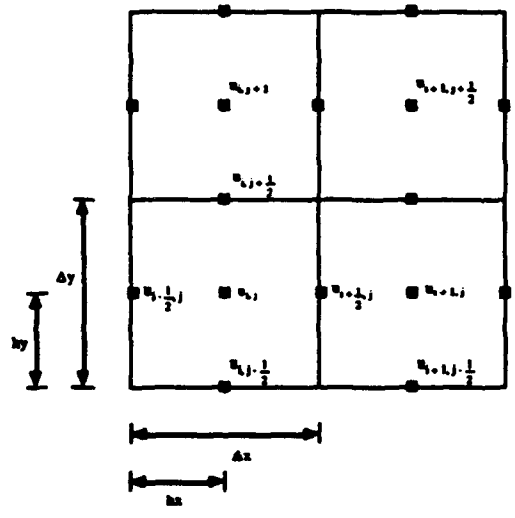


Fig. 2.

Consider the two dimensional diffusion equation $u_t = u_{xx} + u_{yy}$. The compact ADI method is given by equations (8) - (10):

$$\frac{u_{i,j}^{n+\frac{1}{2}} - u_{i,j}^{n}}{\tau} = F_{i,j}^{n+\frac{1}{2}} + G_{i,j}^{n} \tag{8a}$$

$$\frac{u_{i,j}^{n+1} - u_{i,j}^{n+\frac{1}{2}}}{\tau} = F_{i,j}^{n+\frac{1}{2}} + G_{i,j}^{n+1} \tag{8b}$$

where $F_{i,j}$ and $G_{i,j}$ are standard finite difference expressions for u_{xx} and u_{yy} respectively and $\tau = \Delta t/2$. Using the standard finite difference scheme we get the following equations

$$-u_{i-1,j}^{n+\frac{1}{2}} + (\alpha+2)\, u_{i,j}^{n+\frac{1}{2}} - u_{i+1,j}^{n+\frac{1}{2}} = u_{i,j-1}^{n} + (\alpha-2)\, u_{i,j}^{n} - u_{i,j+1}^{n} \tag{9}$$

$$i, j = 1, 3, \ldots, 2M-1, \quad \alpha = \frac{4(\Delta x)^2}{\Delta t}, \quad \Delta x = \Delta y$$

$$-u_{i,j-1}^{n+1} + (\alpha+2)\, u_{i,j}^{n+1} - u_{i,j+1}^{n+1} = u_{i-1,j}^{n+\frac{1}{2}} + (\alpha-2)\, u_{i,j}^{n+\frac{1}{2}} - u_{i+1,j}^{n+\frac{1}{2}} \tag{10}$$

$$i,j = 1,3, \ldots, 2M-1$$

and the continuity conditions are

$$-u_{i-1,j}^{n+\frac{1}{2}} + 2u_{i,j}^{n+\frac{1}{2}} - u_{i+1,j}^{n+\frac{1}{2}} = 0 \tag{9a}$$

$$i, j = 2, 4, \ldots, 2M-2$$

$$-u_{i,j-1}^{n+1} + 2u_{i,j}^{n+1} - u_{i,j+1}^{n+1} = 0 \tag{10a}$$

$$i,j = 2,4, \ldots, 2M-2$$

The two algebraic systems (9), (9a), and (10), (10a) can be solved by the tridiagonal algorithm.

2. Numerical experiments

One of the primary objectives of this project is to vectorization of 1-D and 2-D diffusion problems. We vectorize the 2-D heat equation for the Cray Y-MP using red-black SOR algorithm.

We consider one analytical solutions for the following diffusion equation and compare them with respective

numerical solutions. The results are recorded in tables 1.

The two dimensional diffusion equation is $u_t = u_{xx} + u_{yy} + f(x,y,t)$, $0 < x < 1, 0 < y < 1, t > 0$. The following analytical solutions are considered:

$$u(x,y,t) = 100x(1 - x)y(1-y)e^{-t}, \text{ and } f(x,y,t) = 200(y-y^2 + x-x^2) - 100(x-x^2)(y-y^2e^{-t}$$

All the problems mentioned above are associated with Dirichlet boundary conditions.

The CPT-ADI is compared with FDM-ADI and FDM-SOR. The functional errors in the method is O(h2) as expected and derivative error is of O(h). The result is shown in table 1.

FDM-ADI $\Delta t = \Delta x^2$

No. of Hx=hy cells	No. of iters	CPU Time (sec)	Max Fct Error	Max Derivative Error
16	128	0.76	4.5924E-03	1.01704E-02
32	512	6.15	1.1469E-03	5.24588E-02
64	2048	49.48	2.8667E-04	2.66432E-02

CPT-ADI $\Delta t = \Delta x$

No. of Hx=hy cells	No. of iters	CPU Time (sec)	Max Fct Error	Max Derivative Error
16	16	1.183	5.3563E-03	5.33550E-03
32	32	9.26	1.3391E-03	1.3390E-03
64	64	73.23	3.3479E-03	3.3477E-03

FDM-SOR $\Delta t = \Delta x$

No. of Hx=hy cells	No. of iters	CPU Time (sec)	Max Fct Error	Max Derivative Error
16	240 $\omega = 1.65$	2.10	3.5050E-03	5.869E-01
32	960 $\omega = 1.80$	36.73	2.155E-03	2.940E-01
64	3392 $\omega = 1.85$	473.18	9.5737E-04	1.469E-01

3. Concluding remarks

The x, y domains are devided into 8, 16, 32 and 64 cells and in each case, the starting time is taken to be 0 and the final time is 1. The relation between time step and space step varies among FDM-ADI, FDM-SOR and CPT-ADI schemes. We assume all the material constants are equal to unity and $\Delta x = \Delta y$. In case of FDM-SOR and CPT-ADI, $\Delta x = \Delta y = \Delta t$ and however in case of FDM-ADI, $\Delta t = (\Delta x)^2$. This is the disadvantage for FDM-ADI. For example, when Δx = 1.5625E-2, Δt = 4.883E-4 and it would take 2048 iterations to reach time 1. However, CPT-ADI and FDM-SOR are free from this difficulty. But SOR also has a different disadvantage. Each time step, FDM-SOR takes large number of iterations to converge to a preassigned level. With 64 cells CPT-ADI requires only 64 iterations to reduce the error level to 3.34792E-03, where as FDM-ADI requires 2048 iterations to reduce the error level to 2.86671E-04 and FDM-SOR requires 3,392 iterations to reduce the error level to 9.57374E-04.

As CPT algorithm requires more points evaluation per iteration, it is obvious that CPT requires more CPU time than FDM-SOR or FDM-ADI.

Acknowledgments: Thanks to M. E. Rose for his many valuable suggestions for the completion of this paper. Also many thanks go to A. Nachman of AFOSR.

References

[1] M. E. Rose, "Compact finite volume methods for the diffusion equations," J. Sci. Comp., Vol 4, No. 3 (1989), pp. 261-290.

ATTACHMENT #6 A Comparative Study of Compact Finite Volume Methods for 3-D Diffusion Equation with Finite Volume ADI and SOR

A Comparative Study of Compact Finite Volume Methods For the 3-D Diffusion Equations with Finite Difference ADI and SOR *.

B. N.Borah[1], R.E.White[2], A.Kyrillidis[2], S.Shankarlingam [1], Y.Ji[1]

SSST '93

THE TWENTY-FIFTH SOUTHEASTERN SYMPOSIUM ON SYSTEM THEORY

University of Alabama
Tuscaloosa, Alabama

Sponsored by the
IEEE Computer Society
Technical Committee on Computer Graphics
in cooperation with
ACM/SIGGRAPH

A Comparative Study of Compact Finite Volume Methods For the 3-D Diffusion Equations with Finite Difference ADI and SOR *.

B. N.Borah[1], R.E.White[2], A.Kyrillidis[2], S.Shankarlingam [1], Y.Ji[1]

1. N.C. A & T State University, Greensboro, NC
2. N.C. State University, Raleigh, NC

Abstract

A recently developed Compact Finite Difference Scheme is applied to 3 - D diffusion equations. The relative merits of CPT - ADI is compared with two other computational schemes such as Finite Difference SOR and Finite Difference ADI. The numerical results obtained from these three schemes are compared to known analytical solutions. The primary interest of this paper lies in vectorization and parallel processing. CPT-ADI is the fastest algorithm than both FDM-ADI and FDM -SOR as is evident from the CPU times.

1. Introduction

We shall briefly describe the Compact Finite difference scheme for a 1-D steady state problem. The extension to a 3-D case maybe easily done. The idea behind this approach is to solve the differential equation in isolations from its neighboring subintervals (i.e. compactly) and then extend the solution in the large by means of the continuity conditions for the flux and the temperature across the boundaries of the contiguous subintervals (See Rose[1]).

We consider here 1-D Steady diffusion problem :

(1) D.E. $u' = f$

(2) $\Phi' = u$

(3) B. C. $\Phi = g$

for $X \varepsilon I$ and $I = [I_-, I_+]$.

Divide the interval I into m nonoverlapping subintervals:

$$I_j = \{x / x_{j-1/2} \le x \le x_{j+1/2}\}$$

with center points

$$x_j,\ j = \tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots, M-\tfrac{1}{2}$$

and interior endpoints

$$x_j,\ j = 1, 2, \ldots, M-1$$

We shall adopt finite difference notations

$$\Delta x_i = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}},\ h_i = \frac{\Delta x}{2}$$

and $u(i) = u_i$. We also denote :

$I_c = \left\{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \ldots, M-\tfrac{1}{2}\right\}$ (Center Po ints)

$I_e = \{1, 2, 3, \ldots, M-1\}$ (inte rio r endpoints)

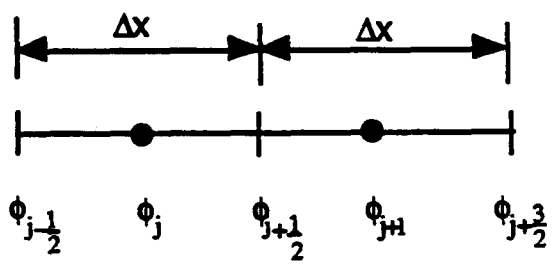


Fig. 1

The discrete equations for (1) and (2) can be written as , for $j \varepsilon I_c$,

(4) $$u_{j+\frac{1}{2}} - u_{j-\frac{1}{2}} = \Delta x_j f_j$$

(5 a) $$\varphi_{j+\frac{1}{2}} - \varphi_j = h_j u_{j+\frac{1}{2}}$$

0094-2898/93 $03.00 © 1993 IEEE

and

$$(5\,b)\qquad \varphi_j - \varphi_{j-\frac{1}{2}} = h_j u_{j-\frac{1}{2}}$$

Next, it is required that both u and f be continuous across every endpoint common to two intervals :

$$[u]_i = [\varphi]_i = 0 \quad \text{for } i \varepsilon I_e$$

Using this continuity condition, $[u]_i = 0$, in terms of φ we get

$$\frac{\varphi_i - \varphi_{i-\frac{1}{2}}}{h_{i-\frac{1}{2}}} = \frac{\varphi_{i+\frac{1}{2}} - \varphi_i}{h_{i+\frac{1}{2}}}$$

and in case of equal intervals, this reduces to

$$(6)\ \varphi_i = \frac{\varphi_{i+\frac{1}{2}} + \varphi_{i-\frac{1}{2}}}{2},\ i \varepsilon I_e \text{ (endpoints)}$$

Now, from (4) and (5), we get

$$(7a)\ \varphi_{j+\frac{1}{2}} - 2\varphi_j + \varphi_{j-\frac{1}{2}} = 2h_j^2 f_j,\ i \varepsilon I_c$$

and from (6) we have

$$(7b)\ \varphi_{j+\frac{1}{2}} - 2\varphi_j + \varphi_{j-\frac{1}{2}} = 0,\ i \varepsilon I_e$$

The equations (7) and the boundary conditions leads to a determined system of algebraic systems for the values of φ, which can be solved by Thomas Algorithm.

The extension to a three dimensional scheme may be done easily.

Three Dimensional stencil is shown in Fig. 2.

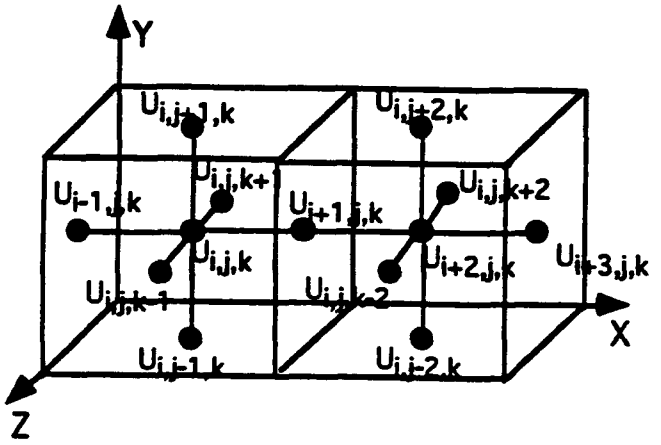


Fig. 2.

Consider the 3-D diffusion equation $u_t = u_{xx} + u_{yy} + u_{zz}$. The Compact ADI method is given by equations (9) - (12) :

$$(8a)\quad \frac{u_{i,j,k}^{n+\frac{1}{4}} - u_{i,j,k}^{n}}{\sigma} = 2F_{i,j,k}^{n} + G_{i,j,k}^{n+\frac{1}{4}}$$

$$(8b)\quad \frac{u_{i,j,k}^{n+\frac{1}{2}} - u_{i,j,k}^{n+\frac{1}{4}}}{\sigma} = G_{i,j,k}^{n+\frac{1}{4}} + 2H_{i,j,k}^{n+\frac{1}{2}}$$

$$(8c)\quad \frac{u_{i,j,k}^{n+\frac{3}{4}} - u_{i,j,k}^{n+\frac{1}{2}}}{\sigma} = G_{i,j,k}^{n+\frac{3}{4}} + 2H_{i,j,k}^{n+\frac{1}{2}}$$

$$(8d)\quad \frac{u_{i,j,k}^{n+1} - u_{i,j,k}^{n+\frac{3}{4}}}{\sigma} = G_{i,j,k}^{n+\frac{3}{4}} + 2F_{i,j,k}^{n+1}$$

where $F_{i,j,k}$, $G_{i,j,k}$ and $H_{i,j,k}$ are finite difference approximation to u_{xx}, u_{yy} and u_{zz} respectively and $\sigma = \frac{\Delta t}{4}$. Using the standard finite difference scheme, we get the folllowing equations :

$$(9)\ -\lambda_y u_{i,j+1,k}^{n+\frac{1}{4}} + (1+2\lambda_y) u_{i,j,k}^{n+\frac{1}{4}} - \lambda_y u_{i,j-1,k}^{n+\frac{1}{4}}$$

$$= 2\lambda_x u_{i+1,j,k}^{n} + (1-4\lambda_x) u_{i,j,k}^{n} + \lambda_x u_{i-1,j,k}^{n}$$

$i,j,k = 1,3, \ldots, 2M-1,\ \lambda_x = \lambda_y = 4.0,\ \Delta x = \Delta y = \Delta z.$

$$(10)\ -2\lambda_z u_{i,j,k+1}^{n+\frac{1}{2}} + (1+4\lambda_z) u_{i,j,k}^{n+\frac{1}{2}} - 2\lambda_z u_{i,j,k-1}^{n+\frac{1}{2}}$$

$$= \lambda_y u_{i,j+1,k}^{n+\frac{1}{4}} + (1-2\lambda_y) u_{i,j,k}^{n+\frac{1}{4}} + \lambda_y u_{i,j-1,k}^{n+\frac{1}{4}}$$

$i,j,k = 1,3, \ldots, 2M-1,\ \lambda_z = \lambda_y = 4.0,\ \Delta x = \Delta y = \Delta z$

$$(11)\ -\lambda_y u_{i,j+1,k}^{n+\frac{3}{4}} + (1+2\lambda_y) u_{i,j,k}^{n+\frac{3}{4}} - \lambda_y u_{i,j-1,k}^{n+\frac{3}{4}}$$

$$= -2\lambda_z u_{i,j,k+1}^{n+\frac{1}{2}} + (1-4\lambda_z) u_{i,j,k}^{n+\frac{1}{2}} - 2\lambda_z u_{i,j,k-1}^{n+\frac{1}{2}}$$

$i,j,k = 1,3, \ldots, 2M-1,\ \lambda_z = \lambda_y = 4.0,\ \Delta x = \Delta y = \Delta z$

$$(12)\quad -2\lambda_x u^{n+1}_{i+1,j,k} + (1+4\lambda_x) u^{n+1}_{i,j,k} - 2\lambda_x u^{n+1}_{i-1,j,k}$$

$$= \lambda_y u^{n+\frac{3}{4}}_{i,j+1,k} + (1-2\lambda_y) u^{n+\frac{3}{4}}_{i,j,k} + \lambda_y u^{n+\frac{3}{4}}_{i,j-1,k}$$

i,j,k=1,3, ...2M-1, $\lambda_x = \lambda_y = 4.0$, $\Delta x = \Delta y = \Delta z$

and the continuity conditions are

$$(9a)\quad -u^{n+\frac{1}{4}}_{i,j+1,k} + 2u^{n+\frac{1}{4}}_{i,j,k} - u^{n+\frac{1}{4}}_{i,j-1,k} = 0$$

$$(10a)\quad -u^{n+\frac{1}{2}}_{i,j,k+1} + 2u^{n+\frac{1}{2}}_{i,j,k} - u^{n+\frac{1}{2}}_{i,j,k-1} = 0$$

$$(11a)\quad -u^{n+\frac{3}{4}}_{i,j+1,k} + 2u^{n+\frac{3}{4}}_{i,j,k} - u^{n+\frac{3}{4}}_{i,j-1,k} = 0$$

$$(12a)\quad -u^{n+1}_{i+1,j,k} + 2u^{n+1}_{i,j,k} - u^{n+1}_{i-1,j,k} = 0$$

i,j,k=2,4, ...2M, $\lambda_x = \lambda_y = 4.0$, $\Delta x = \Delta y = \Delta z$

The four algebraic systems (9),(9a),(10),(10a), (11),(11a) and (12) & (12a) can be solved by the tridiagonal algorithm.

The Finite Difference ADI euations for 3-D diffusion equation are given below:

$$(13)\quad \frac{u^{n+\frac{1}{3}}_{i,j,k} - u^{n}_{i,j,k}}{\sigma} - u^{n+\frac{1}{3}}_{xx} = u^{n}_{yy} + u^{n}_{zz}$$

$$(14)\quad \frac{u^{n+\frac{2}{3}}_{i,j,k} - u^{n+\frac{1}{3}}_{i,j,k}}{\sigma} - u^{n+\frac{2}{3}}_{yy} = u^{n+\frac{1}{3}}_{xx} + u^{n+\frac{1}{3}}_{zz}$$

$$(15)\quad \frac{u^{n+1}_{i,j,k} - u^{n+\frac{2}{3}}_{i,j,k}}{\sigma} - u^{n+1}_{zz} = u^{n+\frac{2}{3}}_{xx} + u^{n+\frac{2}{3}}_{yy}$$

Here $\sigma = \frac{\Delta t}{3}$.

Using the standard Finite Difference scheme we get the following equations :

$$-\lambda_x u^{n+\frac{1}{3}}_{i+1,j,k} + (1+2\lambda_x) u^{n+\frac{1}{3}}_{i,j,k} - \lambda_x u^{n+\frac{1}{3}}_{i-1,j,k}$$

$$(16)\quad = \lambda_y u^{n}_{i,j+1,k} + \lambda_y u^{n}_{i,j-1,k} + \lambda_z u^{n}_{i,j,k+1} + \lambda_z u^{n}_{i,j,k-1} + (1-2\lambda_y - 2\lambda_z) u^{n}_{i,j,k}$$

$$-\lambda_y u^{n+\frac{2}{3}}_{i,j+1,k} + (1+2\lambda_y) u^{n+\frac{2}{3}}_{i,j,k} - \lambda_y u^{n+\frac{2}{3}}_{i,j-1,k}$$

$$(17)\quad = \lambda_x u^{n+\frac{1}{3}}_{i+1,j,k} + \lambda_x u^{n+\frac{1}{3}}_{i-1,j,k} + \lambda_z u^{n+\frac{1}{3}}_{i,j,k+1} + \lambda_z u^{n+\frac{1}{3}}_{i,j,k-1} + (1-2\lambda_x - 2\lambda_z) u^{n+\frac{1}{3}}_{i,j,k}$$

$$-\lambda_z u^{n+1}_{i,j,k+1} + (1+2\lambda_z) u^{n+1}_{i,j,k} - \lambda_z u^{n+1}_{i,j,k-1}$$

$$(18)\quad = \lambda_x u^{n+\frac{2}{3}}_{i+1,j,k} + \lambda_x u^{n+\frac{2}{3}}_{i-1,j,k} + \lambda_y u^{n+\frac{2}{3}}_{i,j+1,k} + \lambda_y u^{n+\frac{2}{3}}_{i,j-1,k} + (1-2\lambda_x - 2\lambda_y) u^{n+\frac{2}{3}}_{i,j,k}$$

Finally , the Finite Difference SOR scheme for the 3-D Diffusion Equation is given below :

$$(1+6\lambda)\text{Utemp} = u^{m}_{i,j,k} + \Delta t f^{m}_{i,j,k} + \lambda(u^{m+1}_{i,j-1,k} + u^{m+1}_{i-1,j,k} + u^{m}_{i+1,j,k} + u^{m}_{i,j+1,k} + u^{m}_{i,j,k-1} + u^{m}_{i,j,k+1})$$

$$u^{m+1}_{i,j,k} = (1-\omega) u^{m}_{i,j,k} + \omega \text{Utemp}, \quad 1 \le \omega \le 2,$$

Here $\lambda = 4.0$ and m is the iteration index.

2. Numerical Experiments

One of the primary objectives of this project is to vectorization of the 3-D problems. We use red-black SOR algorithm to vectorize the 3-D heat equation for the Cray Y-MP.

We consider one analytical solutions for the following diffusion equation and compare them with respective numerical solution. The results are recorded in Table 1.

The 3-D diffusion equation is $u_t = u_{xx} + u_{yy} + u_{zz} + f(x,y,z,t)$, $0 < x < 1$, $0 < y < 1$, $0 < z < 1$, $t > 0$.

The following analytical solutions is considered:
$u(x,y,z,t) = 100x(1-x)y(1-y)z(1-z)e^{-t}$ and
$f(x,y,z,t) = (200((y-y^2)(z-z^2)+(y-y^2)(x-x^2)+(z-z^2)(x-x^2)) - 100x(1-x)y(1-y)z(1-z))e^{-t}$

All the problems mentioned above are associated with Dirichlet boundary conditions.

The CPT-ADI is compared with FDM-ADI and FDM-SOR . The Functional errors in the method is $O(h^2)$ and derivative error is $O(h)$. The result is shown in Table 1.

CPT-ADI

# of cells	iter	CPU Time (sec)	Max Fct Error	Max Der Error
4	4	0.0042	0.0685	0.2828
8	8	0.0912	0.0341	0.1326
16	16	2.3406	0.0046	0.0159

FDM-ADI

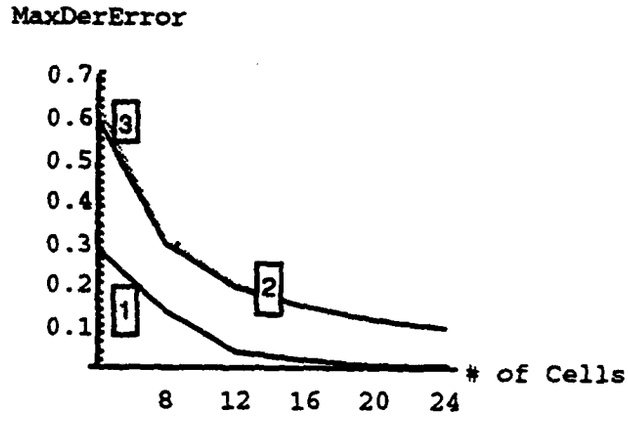
# of cells	iter	CPU Time (sec)	Max Fct Error	Max Der Error
4	16	0.012	0.00977	0.6179
8	64	0.366	0.00241	0.2974
16	256	10.72	0.00006	0.1461

FDM-SOR

# of cells	iter omega	CPU Time (sec)	Max Fct Error	Max Der Error
4	36 1.45	0.0043	0.00318	0.584
8	72 1.55	0.1008	0.00103	0.2909
16	304 1.75	3.158	0.00025	0.1449

Table 1

The curve (1) is for the CPT-ADI scheme, curve (2) is for the FDM-ADI scheme while curve(3) is for the FDM -SOR scheme.



curve (1) above shows the graph of the maxder error Vs. the number of cells for the CPT-ADI scheme. Curve (2) shows the graph of maxder error Vs.number of cells for the FDM-SOR scheme and curve(3) shows the graph of maxder error Vs. the number of cells for the FDM-ADI scheme.

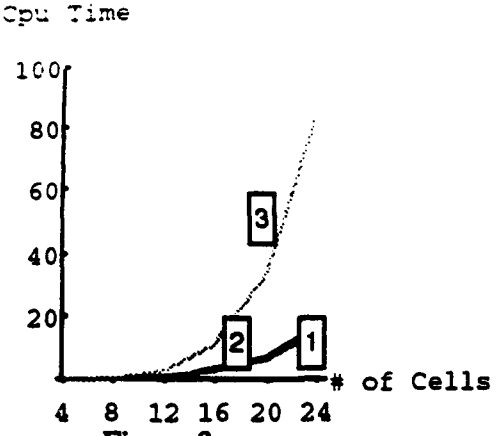


Figure 3

curve (1) is for the CPT - ADI scheme,
curve (2) is for the FDM-SOR scheme and
curve (3) is for the FDM-ADI scheme.

3. Concluding Remarks

The x, y, and z domain are divided into 4, 8 , 16 cells and in each case, the initial time is taken to be 0 and the final time is 1. We assume all the material constants to be equal to unity and $\Delta x = \Delta y = \Delta z$. In case of FDM-SOR and CPT-ADI, $\Delta x = \Delta y = \Delta z = \Delta t$ and however in case of FDM-ADI , $\Delta t = (\Delta x)^2$. This is the disadvantage of FDM-ADI. However , CPT-ADI and FDM-SOR are free from this difficulty. But SOR has a different disadvantage, for each timestep, FDM-SOR takes large number of iterations to converge to a preassigned level. With 16 cells , CPT-ADI requires 16 iterations to reduce the error level to 4.59 E -3 , whereas FDM-ADI requires 256 iterations to reduce the error level to 6.0 E -4 , whereas FDM-SOR requires 304 iterations to reduce the error level to 2.4872 E-4.

In comparing CPU times among all the three approaches, it is found that CPU time for FDM-ADI is the worst. For the problem with 16 cells , CPT-ADI takes only 2.346 sec. However , FDM-ADI approach takes 10.72 secs for the same problem.

In terms of derivative error, CPT-ADI is the best, it is close to the order of $O(h^2)$. The error analysis for the three approaches are shown in Figure 3.

References

[1] Rose, Milton E. Compact Finite Volume Methods for the Diffusion Equation. Journal Scentific Computing, Vol 4, No. 3, Sept. 1989, pp 261 - 290.